

# The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: selecting emission line galaxies using the Fisher discriminant

A. Raichoor<sup>1,\*</sup>, J. Comparat<sup>2,3</sup>, T. Delubac<sup>4</sup>, J.-P. Kneib<sup>4,5</sup>, C. Yèche<sup>1</sup>, H. Zou<sup>6</sup>, F.B. Abdalla<sup>7,8</sup>, K. Dawson<sup>9</sup>, A. de la Macorra<sup>10</sup>, X. Fan<sup>11</sup>, Z. Fan<sup>6</sup>, Z. Jiang<sup>6</sup>, Y. Jing<sup>12</sup>, S. Jouvel<sup>7</sup>, D. Lang<sup>13</sup>, M. Lesser<sup>11</sup>, C. Li<sup>14</sup>, J. Ma<sup>6</sup>, J.A. Newman<sup>15</sup>, J. Nie<sup>6</sup>, N. Palanque-Delabrouille<sup>1</sup>, W.J. Percival<sup>16</sup>, F. Prada<sup>17,18,3</sup>, S. Shen<sup>14</sup>, J. Wang<sup>6</sup>, Z. Wu<sup>6</sup>, T. Zhang<sup>6</sup>, X. Zhou<sup>6</sup>, Z. Zhou<sup>6</sup>

<sup>1</sup> CEA, Centre de Saclay, IRFU/SPP, F-91191 Gif-sur-Yvette, France

<sup>2</sup> Departamento de Física Teórica, Universidad Autónoma de Madrid, Cantoblanco E-28049, Madrid, Spain

<sup>3</sup> Instituto de Física Teórica, (UAM/CSIC), Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain

<sup>4</sup> Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

<sup>5</sup> Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, F-13388, Marseille, France

<sup>6</sup> Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, 100012, China

<sup>7</sup> Department of Physics and Astronomy, University College London, Gower Street, London WC1E6BT, UK

<sup>8</sup> Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa

<sup>9</sup> Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA

<sup>10</sup> Instituto de Física, Universidad Nacional Autónoma de México, A.P. 20-364, 01000, México, D. F. México

<sup>11</sup> Steward Observatory, University of Arizona, Tucson, AZ 85721, USA

<sup>12</sup> Center for Astronomy and Astrophysics, Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>13</sup> McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213, USA

<sup>14</sup> Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

<sup>15</sup> Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA

<sup>16</sup> Institute of Cosmology & Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth, PO1 3FX, UK

<sup>17</sup> Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, E-18080 Granada, Spain

<sup>18</sup> Campus of International Excellence UAM+CSIC, Cantoblanco, E-28049 Madrid, Spain

Preprint online version: January 8, 2016

## ABSTRACT

We present a new selection technique of producing spectroscopic target catalogues for massive spectroscopic surveys for cosmology. This work was conducted in the context of the extended Baryon Oscillation Spectroscopic Survey (eBOSS), which will use  $\sim 200\,000$  emission line galaxies (ELGs) at  $0.6 \leq z_{\text{spec}} \leq 1.0$  to obtain a precise baryon acoustic oscillation measurement. Our proposed selection technique is based on optical and near-infrared broad-band filter photometry. We used a training sample to define a quantity, the Fisher discriminant (linear combination of colours), which correlates best with the desired properties of the target: redshift and [OII] flux. The proposed selections are simply done by applying a cut on magnitudes and this Fisher discriminant. We used public data and dedicated SDSS spectroscopy to quantify the redshift distribution and [OII] flux of our ELG target selections. We demonstrate that two of our selections fulfil the initial eBOSS/ELG redshift requirements: for a target density of  $180\,\text{deg}^{-2}$ ,  $\sim 70\%$  of the selected objects have  $0.6 \leq z_{\text{spec}} \leq 1.0$  and only  $\sim 1\%$  of those galaxies in the range  $0.6 \leq z_{\text{spec}} \leq 1.0$  are expected to have a catastrophic  $z_{\text{spec}}$  estimate. Additionally, the stacked spectra and stacked deep images for those two selections show characteristic features of star-forming galaxies. The proposed approach using the Fisher discriminant could, however, be used to efficiently select other galaxy populations, based on multi-band photometry, providing that spectroscopic information is available. This technique could thus be useful for other future massive spectroscopic surveys such as PFS, DESI, and 4MOST.

**Key words.** Methods: data analysis – Galaxies: general – Galaxies: stellar content – Galaxies: star formation

## 1. Introduction

Large optical imaging surveys in astronomy, such as the Sloan Digital Sky Survey (SDSS, York et al. 2000) or the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS, Gwyn 2012), have revolutionised the fields of galaxy evolution and cosmology. Indeed, they enable the photometric selection of large,

controlled galaxy populations over either a very wide area or up to faint magnitudes. Such large, homogeneous galaxy samples are needed to define target catalogues for intensive spectroscopic surveys (e.g., the Main Galaxy Sample: Strauss et al. 2002; the VIMOS VLT Deep Survey, VVDS: Le Fèvre et al. 2005, 2013). Those spectroscopic surveys are then used to measure galaxy properties (e.g., Kauffmann et al. 2003; Ilbert et al. 2005) or cosmological parameters (e.g., Eisenstein et al. 2005; Percival et al.

\* e-mail: anand.raichoor@cea.fr

2010; Anderson et al. 2012) with high statistical accuracy. In addition, the photometric galaxy samples themselves can also put interesting constraints on galaxy evolution or cosmology (e.g., van Dokkum et al. 2010; Seo et al. 2012).

In this context, a significant step was taken by the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2013), which uses 1.5 million galaxies over  $10\,000\text{ deg}^2$  selected in the SDSS, to precisely measure the scale of the baryon acoustic oscillations (BAO) to redshifts  $z < 0.6$ , and 160 000 quasars to produce measurements at  $z > 2.1$  using the quasar Lyman- $\alpha$  forest (Delubac et al. 2014; Font-Ribera et al. 2014).

The main goal of the BOSS survey is to put a cosmological constraint on dark energy through the measurement of the BAO, but its legacy for galaxy evolution will also be unique. The BOSS observations ended in early 2014, and the data were released as part of the DR11 and DR12 (Alam et al. 2015). Building on the success of the BOSS survey, the *extended Baryon Oscillation Spectroscopic Survey* (eBOSS, Dawson et al., 2015) will use four different tracers of the underlying density field to expand the volume covered by BOSS focussing on the redshift range  $0.6 < z < 2.2$ . The four eBOSS tracers are *i*) luminous red galaxies (LRGs) at  $z \sim 0.7$ , *ii*) emission line galaxies (ELGs) at  $z \sim 0.8$ , *iii*) quasars (at  $0.9 < z < 2.2$ ), and *iv*) Lyman- $\alpha$  absorbers in the line of sight of high-redshift ( $2.1 < z < 3.5$ ) quasars.

For the ELG tracers, 300 spectroscopic plates (BOSS plates have a  $\sim 7\text{ deg}^2$  circular area and 900 fibres) are to be dedicated to observing 270 000 targets as potential ELGs in the South Galactic Cap (SGC). The choice of targeting ELGs is motivated by the presence of the [OII] emission line in the ELG spectrum, which permits an efficient redshift measurement with  $\sim 1\text{h}$  of exposure time. The requirement for ensuring a measurement of BAO parameters with a precision of 2% with ELGs is to obtain the spectroscopic redshift measurements of 190 000 ELGs in the  $0.6 < z_{\text{spec}} < 1.0$  redshift range with a precision better than  $300\text{ km.s}^{-1}$ , with  $\lesssim 1\%$  of catastrophic failures (precision greater than  $1000\text{ km.s}^{-1}$ ). To fulfil those requirements, the initial eBOSS/ELG settings are an area of  $1500\text{ deg}^2$  and a  $180\text{ deg}^{-2}$  target density with a minimal efficiency of 70%, where we define the efficiency as the number of ELGs with a reliable  $z_{\text{spec}}$  measurement with  $0.6 < z_{\text{spec}} < 1.0$  divided by the number of targets. We used those values for the baseline of this study. The technique we propose here has the advantage of being flexible, thus could be adapted in the case where the final eBOSS/ELG requirements should differ from those values. For instance, for the eBOSS ELG programme a fibre density of  $170\text{ deg}^{-2}$  is assumed in Dawson et al. (2015) – hence shifting the minimum required efficiency to 74% – as  $10\text{ deg}^{-2}$  fibres are reserved for other targets.

This paper is part of a series of papers analysing the properties of  $z \sim 0.8$  ELG selection, paving the way for the final eBOSS ELG selection. This paper (Paper II) introduces a new method of selecting  $z \sim 0.8$  ELG based on the SDSS detected objects and describes the redshift and [OII] properties of the selected galaxies. In Delubac et al. (2015: Paper III), we present the catalogue of the selected ELGs, along with various homogeneity and systematics tests. Comparat et al. (2015a, Paper I) study the [OII], H $\beta$ , and [OIII] emission lines measurement at  $z \sim 0.8$  with the BOSS spectrograph (Smee et al. 2013), aiming to better understand the redshift estimation and the selected galaxy properties. It also details the spectroscopic observations dedicated to the preliminary study of ELG selection. Jouvel et al. (2015, Paper IV) analyses the properties (redshift, homogeneity) of a  $z \sim 0.9$  ELG selection based on the Dark Energy Survey

(DES<sup>1</sup>; The Dark Energy Survey Collaboration 2005) photometry.

In this paper, we present a novel method of select  $z \sim 0.8$  ELGs. Compared to the initial tests, which only used the optical bands (see Paper I), our analysis additionally includes one near-infrared band, hence adding one dimension to the colour-colour space. The most common method of selecting galaxies for a spectroscopic survey is to apply cuts in magnitudes and colour-colour spaces. It has been used for the surveys targeting a given redshift range (e.g., DEEP2: Newman et al. 2013; VIPERS: Guzzo et al. 2014) or for the surveys used for BAO measurements (e.g., SDSS/LRG: Eisenstein et al. 2001; WiggleZ: Drinkwater et al. 2010; the upcoming DESI/LRG-ELG<sup>2</sup>). However, such an approach has some limitation when using a large number of multi-wavelength observations: when dealing with three or more colour-colour diagrams, the selection-box definition starts to be subjective, unless using an automatic exploration of all the possibilities. One possibility is to use neural networks (e.g., for the BOSS/QSOs: Dawson et al. 2013), which can bring efficient selections but at the cost of a less tractable selection. We introduce an alternative approach, the Fisher discriminant, which is equivalent to a hyperplane cut in the full colour space, that is to say, a cut on a simple linear combination of the colours. This hyperplane is automatically defined from a training sample and a list of criteria, which are here a redshift of  $\sim 0.8$  and significant [OII] emission. We note that this approach – not used in astrophysics to our knowledge – is automatic and can be used in other situations where one wants to select a given population from multi-wavelength photometry, given that a training sample is available. We present in Section 2 the Fisher discriminant approach, then we introduce in Section 3 the photometric and spectroscopic data used in this study. Section 5 is dedicated to the tested  $z \sim 0.8$  ELG selection schemes: we first describe them and then analyse their global properties in terms of redshift and [OII] emission. For two of the selections, we present stacked spectra and structural properties in Section 6. Finally, we conclude in Section 7.

In this paper, we adopt  $H_0 = 70\text{ km s}^{-1}\text{ Mpc}^{-1}$ ,  $\Omega_m = 0.30$ , and  $\Omega_\Lambda = 0.70$ . All magnitudes are expressed in the AB system and corrected for the Galactic foreground extinction using the Schlegel et al. (1998) maps.

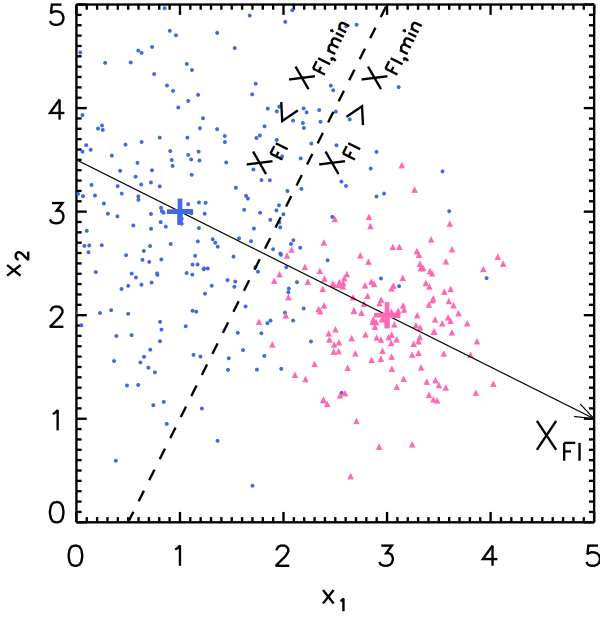
## 2. Fisher discriminant method

### 2.1. Principle

The goal of the linear discriminant analysis (LDA, Fisher 1936), also known as the Fisher method, is to define a discriminant (the Fisher discriminant  $X_{FI}$ ) that separates two known classes of a set of events best. We assume we have a collection of events ( $\mathbf{y}$ ) where each event  $\mathbf{y}$  is known to belong to one of the two classes, ( $\mathbf{y}_1$ ) and ( $\mathbf{y}_2$ ). To each event are associated  $N$  measurements:  $\mathbf{y} = (x_1, x_2, \dots, x_N)$ , each  $x_i$  being a real variable measuring a given property. For instance, in the original taxonomic work of Fisher (1936), the two classes ( $\mathbf{y}_1$ ) and ( $\mathbf{y}_2$ ) are two different species of iris, *Iris setosa* and *Iris versicolor*; the events  $\mathbf{y} = (x_1, x_2, x_3, x_4)$  are a sample of fifty plants, for each of which are available four measurements done on the sepals and the petals. We then let  $n_1$  and  $n_2$  be the number of events in each class,  $\bar{\mathbf{y}}_1$  and  $\bar{\mathbf{y}}_2$  the means over each class, and  $T$  the total variance-covariance matrix of the sample ( $\mathbf{y}$ ).

<sup>1</sup> <http://www.darkenergysurvey.org>

<sup>2</sup> Dark Energy Spectroscopic Instrument: <http://desi.lbl.gov/cdr/>



**Fig. 1.** Illustration of the Fisher discriminant method with  $N = 2$ . The  $(y_1)$  class is in blue dots, with the blue cross at  $\bar{y}_1 = (1, 3)$ , while the  $(y_2)$  class is in pink triangles, with the pink cross at  $\bar{y}_2 = (3, 2)$ . For each event, the Fisher discriminant  $X_{FI}$  corresponds to its orthogonal projection along the axis defined by  $\bar{y}_1$  and  $\bar{y}_2$ . The dashed line illustrates the hyperplane used to split the events in two classes.

The Fisher discriminant  $X_{FI}$  is a linear combination of the  $N$  variables  $x_i$ , aiming to provide the best separation between the two classes of events  $(y_1)$  and  $(y_2)$ . In the  $N$ -dimension space of the measurement variables, it defines a hyperplane (dimension  $N - 1$ ). This hyperplane is orthogonal to the axis defined by the line connecting  $\bar{y}_1$  and  $\bar{y}_2$ , along which the distance between the projected points will naturally be a maximum. In other terms, the Fisher discriminant  $X_{FI}$  is the orthogonal projection on this axis, as illustrated in Figure 1 for  $N = 2$ . In his original work, Fisher (1936) proposed to normalise the projected distance by the quadratic sum of the projected dispersion of each class:

$$X_{FI} = \frac{\sqrt{n_1 n_2}}{n_1 + n_2} (\bar{y}_1 - \bar{y}_2)^T T^{-1} \mathbf{y}. \quad (1)$$

A threshold value  $X_{FI, \min}$  is then used to associate the events with  $X_{FI} < X_{FI, \min}$  to the  $(y_1)$  class and the events with  $X_{FI} > X_{FI, \min}$  to the  $(y_2)$  class.

## 2.2. Application to ELGs

We now describe how we apply the Fisher discriminant approach to ELGs. The considered variable space is the galaxy colour space, and we compute the Fisher discriminant quantity  $X_{FI}$  through the use of a spectroscopic sample of galaxies (presented in Section 3.2). As for any learning method, the training and test sample should ideally be as representative as possible of the data we want to apply the method to. This is the case in our study, where the training and test samples have homogeneous photometry coming from the surveys planned to be eventually used in eBOSS/ELG (except for the DECaLS data that we mimic with degrading the CFHTLS data; see Section 3.1). Please note

that, as we test several selection schemes defined with different Fisher discriminants  $X_{FI}$ , we rescale the Fisher discriminant  $X_{FI}$  obtained with Eq.(1) to a common scale, in order to facilitate the comparison between Fisher expressions. More precisely, we multiply it by a coefficient and add a constant value, so that the  $X_{FI}$  distribution for a complete subsample (the VVDS, see Section 3.2) has a mean value of 0 and a standard deviation of 1.

Our aim is to select ELGs at  $0.6 \leq z_{\text{spec}} \leq 1.0$ : we need to choose how to define the two considered classes, the Signal and the Background classes. A first possible approach to define the two considered classes is to use criteria on both the spectroscopic redshift  $z_{\text{spec}}$  and the total [OII] flux  $f_{\text{[OII]}}^{\text{tot}}$ , with defining our Signal class with galaxies in a high redshift range and with a significant total [OII] flux.

However,  $f_{\text{[OII]}}^{\text{tot}}$  is available for relatively few galaxies compared to  $z_{\text{spec}}$ , which is a quantity usually provided in public surveys. To cover the maximum variable space, we can also take advantage of the fact that in our targeted redshift range, a large majority of galaxies are star-forming (e.g. Ilbert et al. 2013), as the star-formation rate density of the Universe is about ten times higher at  $z \sim 1$  than today (e.g. Lilly et al. 1996; Madau et al. 1998; Hopkins & Beacom 2006). A possible alternate approach is to define our Signal class only with galaxies in a high-redshift range. Proceeding in this way would allow us to use a large spectroscopic training sample, representative of the main types of galaxies at all redshifts.

## 3. Data

In this section, we describe the data used in our study. First, we present the photometric data with which the ELG selection is done, then we describe the Fisher training sample used to define the tested Fisher discriminants  $X_{FI}$ ; finally, we introduce the spectroscopic data used to quantify the efficiencies of the tested Fisher selection algorithms. We display in Figure 2 the sky locations of the different surveys of interest in this study.

### 3.1. Photometry

We here present the different photometric surveys we use in this study to estimate object colours. We test different schemes, based on objects detected in the SDSS. The general properties of the used photometry are summed up in Table 1. We note that the issues related to colours computed with magnitudes measured through different surveys are significantly mitigated for our SDSS-SCUSS-WISE colours, because our SCUSS and WISE photometry is done consistently with the SDSS (forced photometry on SDSS objects, using SDSS structural information: Lang et al. 2014, Hu Zou et al., 2015, in preparation).

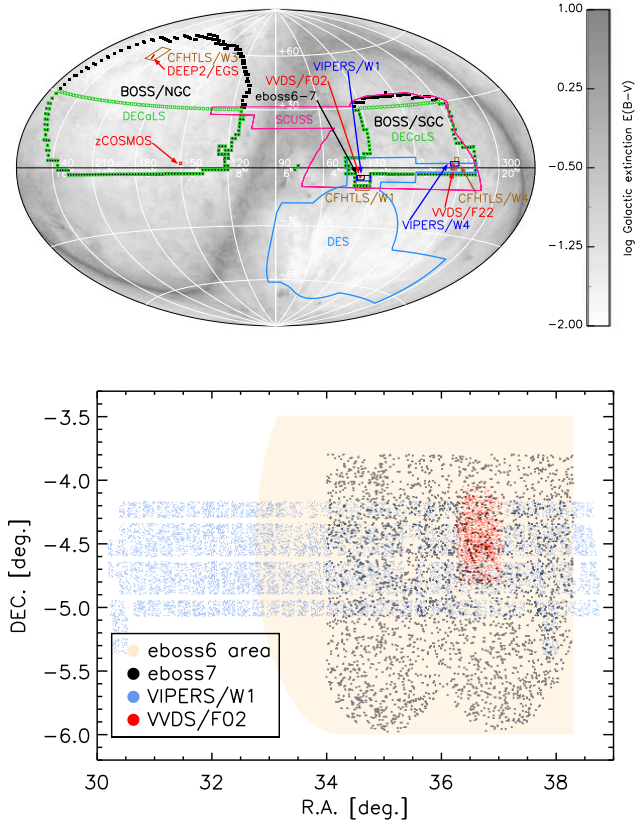
**SDSS.** The SDSS (York et al. 2000; Alam et al. 2015, ; DR12) provides photometry in the optical  $ugriz$  broad-bands to a depth of  $r \sim 22.5$  mag over  $\sim 15\,000$  deg<sup>2</sup> of high-latitude sky, split into two regions of  $\sim 7\,500$  deg<sup>2</sup>, namely the Northern Galactic Cap (NGC) and the Southern Galactic Cap (SGC).

Since we are interested in galaxy colours, we use the modelMag magnitudes<sup>3</sup>: those are computed through a luminosity profile – fitted to the  $r$ -band data – convolved with the PSF in each band, and this permits unbiased colour measurements in the absence of colour gradients. Our photometric object detection list is constituted of objects from the PhotoPrimary list<sup>4</sup>.

<sup>3</sup> [http://www.sdss.org/dr12/algorithms/magnitudes/#mag\\_model](http://www.sdss.org/dr12/algorithms/magnitudes/#mag_model)

<sup>4</sup> <http://skyserver.sdss.org/dr12/en/help/browser/browser.aspx>





**Fig. 2.** Sky location of the different photometric and spectroscopic surveys used in this study. *Top panel:* overview of the whole sky in an Aitoff projection in J2000 equatorial coordinates; the WISE photometry covers the whole sky. *Bottom panel:* zoom on the location of eboss6–7, the ten eBOSS/ELG test plates observed in 2014; in this study, we do not use eboss6 measurements outside of the eboss7 area.

**WISE.** The Wide-Field Infrared Survey Explorer (WISE; Wright et al. 2010) measured the full sky in four mid-infrared bands centred on  $3.4\ \mu\text{m}$ ,  $4.6\ \mu\text{m}$ ,  $12\ \mu\text{m}$ , and  $22\ \mu\text{m}$ , known as W1 through W4. We make use of the “forced photometry” done on the SDSS objects here (Lang et al. 2014): this photometry uses measured SDSS source positions, star-galaxy separation, and galaxy profiles to define the sources whose fluxes are to be measured in the WISE images. In this work, we use only the W1 channel, because it appeared during our preliminary studies that also using W2 creates spatial inhomogeneities in our selections, thus reflecting the underlying variations in the signal-to-noise ratio in W2.

**SCUSS.** The South Galactic Cap  $u$ -band Sky Survey (SCUSS; Xu Zhou et al., 2015, in preparation; Zou et al. 2015) is an international cooperative project, which is undertaken by the National Astronomical Observatories of China (Chinese Academy of Sciences) and Steward Observatory (University of Arizona, USA). It is a  $u$ -band (effective wavelength  $\sim 3538\ \text{\AA}$ ) imaging survey programme with the 90-inch (2.3 m) Bok telescope located on Kitt Peak. This survey has imaged  $\sim 5\,000\ \text{deg}^2$  in the SGC with 80% of the area overlapped with the southern SDSS data, but 1–1.5 magnitudes deeper than the SDSS  $u$ -band photometry. The  $u$ -band filter used in the SCUSS project is similar to the SDSS  $u$ -band filter but slightly bluer. For our aims, this

deep  $u$ -band imaging brings valuable improvement with respect to the SDSS  $u$ -band photometry, given that the  $u$ -band magnitude tightly correlates with the [OII] flux (Comparat et al. 2015b) and that the typical ELGs at  $0.6 \leq z_{\text{spec}} \leq 1.0$  are too faint to have a robust measurement in the SDSS  $u$  band.

For SCUSS we also use a forced *model* photometry on SDSS objects (Hu Zou et al., 2015, in preparation). This photometry constructs 2D models (de Vaucouleurs and exponential) based on SDSS  $r$ -band galaxy profiles and star-galaxy separation, and estimates object fluxes through comparing the models with the object images of SCUSS. The *modelMag* magnitudes in SCUSS is derived from the object flux with higher likelihood in the de Vaucouleurs and exponential model fitting.

**CFHTLS degraded to DECaLS.** We also test a scheme where the SDSS  $z$  band is replaced by the DECaLS  $z$  band. The DECam Legacy Survey<sup>5</sup> (DECaLS) will conduct a three-band imaging survey of the SDSS extragalactic footprint. The Dark Energy Camera (DECam) will be used to image the  $6700\ \text{deg}^2$  footprint overlapping SDSS in the region  $-20 < \text{Dec. [deg]} < 30$ , to depths of  $g = 24.7\ \text{mag}$ ,  $r = 23.9\ \text{mag}$ , and  $z = 23.0\ \text{mag}$ . The survey will be conducted from 2014 through 2017, with periodic data releases beginning in March 2015. As of January 2015,  $>1000\ \text{deg}^2$  have already been observed in the SGC in the  $z$  band to a depth of  $z = 22.7\ \text{mag}$  (whereas the  $g$ - and  $r$ -band observations have a significant lower coverage). To simulate the DECaLS photometry, which is not public yet, we degrade the CFHTLS-Wide W1<sup>6</sup> (Gwyn 2012)  $z$ -band photometry to the depth of the DECaLS  $z$  band photometry by adding random Gaussian noise. For both surveys we do a linear fit to the data:  $\log_{10}(em) = zp + s \times m$ , where  $m$  is the magnitude and  $em$  the magnitude error. Then, if an object has  $m_{\text{CFHTLS}} = m_0$ , we degrade it to  $m_{0,\text{deg}} = m_0 + r \times \sqrt{10^{2(z_{\text{PDECaLS}} + s_{\text{DECaLS}} \times m_0)} - 10^{2(z_{\text{PCFHTLS}} + s_{\text{CFHTLS}} \times m_0)}}$ , where  $r$  is randomly drawn from a Gaussian distribution centred on 0 with a width of 1 and associate to this object a magnitude error of  $10^{z_{\text{PDECaLS}} + s_{\text{DECaLS}} \times m_{0,\text{deg}}}$ . We note that we did not model the scatter in the magnitude error, since this feature has no influence on our selection process.

In addition to the DECaLS, the DES – started in Autumn 2013 – also uses the DECam instrument to image  $5000\ \text{deg}^2$  in the *grizY* bands to about two magnitudes deeper than the DECaLS. Therefore, the DES  $z$ -band photometry for the  $\sim 500\ \text{deg}^2$  overlapping region between the SDSS and DES could also be considered.

### 3.2. Fisher training sample

For each selection scheme, we need to define the Fisher discriminant quantity  $X_{FI}$ , i.e. the linear colour combination. This is done through the use of a spectroscopic training sample, for which the used photometry and the quantity used to define the classes, the spectroscopic redshift  $z_{\text{spec}}$  (and eventually the total [OII] flux  $f_{\text{[OII]}}^{\text{tot}}$ ), are known. As explained in Section 2.2, our approach is to use a large, composite sample of galaxies, in order to cover the loci as much as possible in the colour space occupied by the different type of galaxies at all redshifts. We list below the surveys we use to define our Fisher training sample.

**VIPERS.** The VIMOS Public Extragalactic Redshift Survey (VIPERS, Guzzo et al. 2014) is an on-going large programme

<sup>5</sup> <http://portal.nersc.gov/deca/s/>

<sup>6</sup> <http://terapix.iap.fr/cpl/T0007/doc/T0007-doc.html>

Survey	BOSS/SGC ( $\sim 3,100 \text{ deg}^2$ ) coverage	BOSS/NGC ( $\sim 7,500 \text{ deg}^2$ ) coverage	Band depth (mag)
SDSS	100%	100%	$g = 23.1, r = 22.7, i = 22.2, z = 20.7$
WISE (forced photometry)	100%	100%	$W1 = 20.3$
SCUSS	100%	-	$u \sim 23$
DECaLS/DR1 (CFHTLS-Wide degraded)	$\sim 30\%$	$\sim 45\%$	$z = 22.7$

**Table 1.** Photometric data properties summary (Section 3.1). Reported depths correspond to a  $5\sigma$  point-source detection.

that builds a spectroscopic sample of  $10^5$  galaxies with  $17.5 \leq i \leq 22.5$  and  $0.5 < z_{\text{spec}} < 1.5$  over a total area of  $24 \text{ deg}^2$  within the CFHTLS-Wide W1 and W4 fields. The observations are done with the VIMOS instrument (Le Fèvre et al. 2003) with the LR-RED grism (wavelength coverage:  $\sim 5500\text{--}9500 \text{ \AA}$ ; spectral resolution:  $R \sim 250$ ; 0.75h exposure time). The low-redshift ( $z_{\text{spec}} < 0.5$ ) galaxies are efficiently removed from the target selection through a colour cut, resulting in a completeness  $> 95\%$  in the  $0.6 < z_{\text{spec}} < 1.2$  range (Guzzo et al. 2014). This sample is crucial for our study, because it covers our targeted redshift range and is flux-selected with an  $i$ -band flux fainter than the SDSS  $i$ -band depth. Among the 57 204 spectra of the First Data Release (Garilli et al. 2014; Franzetti et al. 2014), we restricted ourselves to those that have a secure redshift flag ( $2 \leq \text{Flag} < 5$ ) and that are detected in the SDSS.

**Other public surveys.** In our Fisher training sample, we include SDSS galaxies belonging to the following public surveys: the F02 and F22 fields of the VIMOS VLT Deep Survey (VVDS, Le Fèvre et al. 2005, 2013; VIMOS LR-RED grism:  $\sim 5500\text{--}9500 \text{ \AA}$ ;  $R \sim 250$ ; 0.75h and 4.5h), the zCOSMOS 10k-Bright Spectroscopic Sample (Lilly et al. 2009; VIMOS MR grism:  $5500\text{--}9650 \text{ \AA}$ ;  $R \sim 600$ ; 1h exposure time), and the EGS field of DEEP2 (Newman et al. 2013; DEIMOS spectrograph:  $6500\text{--}9300 \text{ \AA}$ ;  $R \sim 5900$ ; 1h exposure time). We consider only objects having reliable spectroscopic redshifts. Those surveys, including VIPERS, are magnitude-limited, with a magnitude limit fainter than the SDSS depth (VVDS/F02:  $17.5 < i < 24$ , VVDS/F22:  $17.5 < i < 22.5$ , zCOSMOS/Bright:  $i < 22.5$ , DEEP2/EGS:  $R < 24.1$ ), so that they include all possible types of galaxies detected in the SDSS (lying in the observed sky region). We also add all the SDSS DR12 public spectroscopic redshifts (DR12, Alam et al. 2015; SDSS and BOSS spectrographs:  $3800\text{--}9200 \text{ \AA}$  and  $3650\text{--}10,400 \text{ \AA}$ ;  $R \sim 1500\text{--}2500$ ; 0.75-1h exposure time) covering those survey regions.

**Comparat et al. (2015) ELGs.** Furthermore, we enlarge our Fisher training sample by adding  $\sim 10^4$  ELGs observed as pilot programmes for eBOSS and DESI (Paper I).

**Total [OII] flux.** In addition, we extracted the total [OII] fluxes for the VIPERS, VVDS, and BOSS/eBOSS surveys in a consistent way from the spectra  $f_{[\text{OII}]}^{\text{tot}}$  (see Paper I for details).

### 3.3. Spectroscopic verification sample

We tested the efficiency of our ELG Fisher selection algorithms in a  $8.82 \text{ deg}^2$  area centred approximately at (R.A., DEC.)  $\sim (36.0, -4.8)$  (see bottom panel of Figure 2), thus using a verification sample that is independent of our training sample ( $\sim 6\%$  overlap). This part of the sky has been extensively observed in 2014 with ten eBOSS/ELG test plates (eboss6: plates 8123-8130; eboss7: plates 8355, 8356), with  $4 \times 15$  minutes expo-

sure using the BOSS spectrograph ( $3650\text{--}10,400 \text{ \AA}$ ;  $R \sim 1500\text{--}2500$ ). The eboss7 plates have been specifically dedicated to two of our tested photometric Fisher discriminant selections (Fisher\_UGRIZW1 and Fisher\_GRIW1, see Section 5), which thus have a spectroscopic coverage of the target selection of  $\sim 93\%$ , the remaining  $\sim 7\%$  having not been targeted for tiling reasons. The observations and their reduction are described in Paper I, to which we refer the reader. In short, the observations were done with the BOSS spectrograph (Smee et al. 2013) of the 2.5m telescope located at Apache Point Observatory (New Mexico, USA), using  $2''$  diameter fibers and an exposure time of  $\sim 1\text{h}$ . The reduction provides various information, with a confidence flag based on the continuum (zCont) and emission lines (zQ). The zCont flag quantifies the degree to which the continuum is detected, and the zQ flag quantifies the number of detected emission lines, along with the signal-to-noise ratio ( $S/N$ ) thereof. We refer to Paper I for further details on the definition of zCont and zQ. We have restricted this study to galaxies having a secure confidence flag; that is,

$$zQ \geq 1.5 \quad (2a)$$

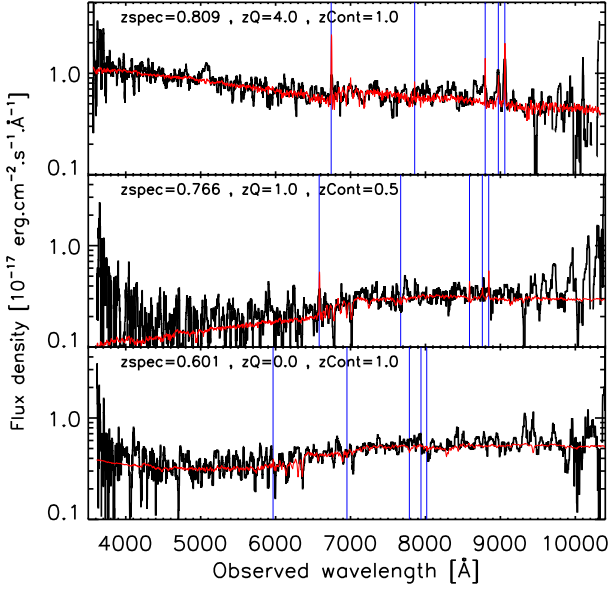
$$\text{or } (zQ=1 \text{ and } zCont \geq 0.5) \quad (2b)$$

$$\text{or } (zQ=0 \text{ and } zCont \geq 1.0), \quad (2c)$$

meaning that the galaxy has either undoubted emission features – one line at  $S/N \geq 5$  or two or more lines at  $S/N \geq 3$  (Eq.(2a)) – or a trustable combination of a detection of emission features and of the continuum: one line at  $S/N \geq 3$  and a continuum detected at  $S/N \geq 8$  with at least three emission lines (Eq.(2b)) or a continuum detected at  $S/N \geq 10$  with at least three emission lines (Eq.(2c)). We illustrate in Figure 3 those flags with three eboss6–7 spectra. The expected catastrophic  $z_{\text{spec}}$  estimate rate, estimated through the visual inspection of more than 10 000 BOSS spectra, can be estimated for each {zCont, zQ}. We note that in the  $0.6 \leq z \leq 1.0$  range, the galaxies selected with Eqs.(2b,2c) have, on average, lower precision in the  $z_{\text{spec}}$  estimate (median value of  $z_{\text{spec,err}}/(1 + z_{\text{spec}})$  of  $1.5\text{--}1.7 \times 10^{-4}$  vs.  $0.5 \times 10^{-4}$ ) and a slightly higher catastrophic  $z_{\text{spec}}$  estimate rate when compared to the galaxies selected with Eq.(2a).

The eboss6–7 data make it possible an unbiased and complete analysis for two of our tested selections (Fisher\_UGRIZW1 and Fisher\_GRIW1). For the remaining tested selections, those eboss6–7 data represent a biased and incomplete subsample of the selections, so they cannot be used to reliably infer the selection properties. To overcome this, we duplicate the analysis using complementary data for all the selections: the CFHTLS-Wide photometric redshifts for the redshift, and the VIPERS total [OII] fluxes for the [OII] diagnosis. The CFHTLS-Wide photometric redshifts (T0007 release<sup>7</sup>; Ilbert et al. 2006; Coupon et al. 2009) are of very good quality up to  $i < 22.5$  (bias below 1%, scatter of  $\sim 0.04$ , and less than 4% outliers). Using the eboss6–7 test plates, we demonstrate the reliability of those photometric redshifts for ELGs up to redshift  $\sim 1$  in Appendix

<sup>7</sup> [http://terapix.iap.fr/rubrique.php?id\\_rubrique=267](http://terapix.iap.fr/rubrique.php?id_rubrique=267)



**Fig. 3.** Example of three eboss6–7 spectra, illustrating the confidence flags ( $zQ$ ,  $zCont$ ). The observed spectra (black) is smoothed: each pixel is replaced by the median value of the 20 nearest pixels. The best-fit model is in red. Vertical blue lines illustrate the location of some expected emission lines in the observed frame:  $[OII]_{3725-3727}$ ,  $H\gamma_{4342}$ ,  $H\beta_{4862}$ ,  $[OIII]_{4959}$ ,  $[OIII]_{5007}$  (in increasing wavelength).

A. As explained in the previous section, the VIPERS sample is magnitude-complete down to  $i = 22.5$  mag in the redshift range  $0.6 \leq z_{\text{spec}} \leq 1.2$ .

Objects with a signal-to-noise ratio ( $S/N$ ) of 3 in the  $[OII]$  flux measurements have on average  $f_{[OII]}^{\text{tot}} \sim 10^{-16.6} \text{ erg.s}^{-1}.\text{cm}^{-2}$  in the VIPERS observations and  $f_{[OII]}^{\text{tot}} \sim 10^{-16.4} \text{ erg.s}^{-1}.\text{cm}^{-2}$  in the eboss6–7 observations.

#### 4. Improvement with using the WISE/W1 data

The WISE/W1 near-infrared data bring crucial information for identifying the galaxy redshift, and the combination with other colours permits isolating  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs. We illustrate this point with the  $r - W1$  vs.  $g - r$  diagram in Figure 4, using both model predictions and data.

In the left-hand panel, we plot the tracks predicted by the Bruzual & Charlot (2003) stellar population models with standard settings ( $z_{\text{form}} = 3$ , solar metallicity) for four different exponentially declining star formation histories (SFH  $\tau = 0.05, 1, 5, 10$  Gyr) with no dust. Models with SFH  $\tau = 5, 10$  Gyr are representative of ELGs, while models with SFH  $\tau = 0.05, 1$  Gyr are representative of LRGs. Besides, it is known that galaxies at  $0.6 \leq z_{\text{spec}} \leq 1.0$  with star formation can be dusty. We represent the effect of a  $E(B - V) = 0.2$  reddening using the Calzetti et al. (2000) law, and this value corresponds to the median value for  $0.6 \leq z \leq 1.0$  star-forming galaxies with  $i < 22.5$  mag in the COSMOS catalogue (Ilbert et al. 2009). To guide the eye, we shade the approximate locus where  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs are expected to lie according to those models.

The assumptions made to compute the models (formation redshift, metallicity, SFH, dust) are simple and generic: to verify that the model predictions agree with observed galaxy proper-

ties, we look at the same colour-colour diagram, but with observed data. We plot the loci of the SDSS objects belonging to complete spectroscopic surveys (see Section 3.2) with  $i < 22.5$  mag. In the middle panel, we gather the VVDS/F22 galaxies in bins of spectroscopic redshift: the colour evolution with the redshift agrees with the model predictions, with  $0.6 \leq z_{\text{spec}} \leq 1.0$  galaxies lying at  $1 \lesssim r - W1 \lesssim 3$  and spanning a wide range of  $g - r$  colours. Then we plot in the right-hand panel the  $0.6 \leq z_{\text{spec}} \leq 1$  VIPERS galaxies binned by  $[OII]$  luminosity. Again, the data and the model predictions agree, with  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs having blue  $g - r \lesssim 1$  colours.

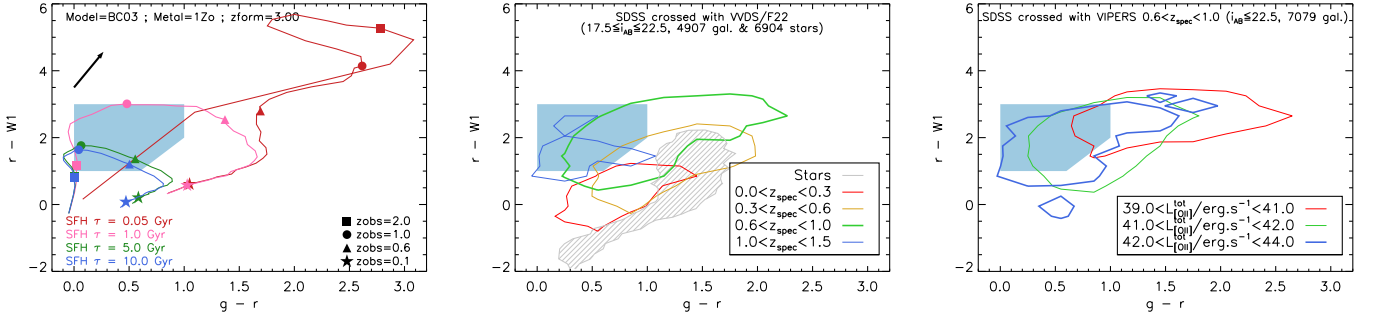
Thus, we see that the  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs can be isolated using colours that include the WISE/W1 data. The  $z_{\text{spec}} < 0.6$  galaxies should lie in a different locus at bluer  $r - W1$  colours; stars are also expected to minimally contaminate the ELG selection (Figure 4, middle panel). The  $z_{\text{spec}} > 1$  galaxies lie at a locus overlapping the blue-shaded area; however, we note that our requirement that the eBOSS/ELG galaxies are detected in the SDSS significantly reduces the contamination from them, because the SDSS is too shallow to detect a large number of  $z_{\text{spec}} > 1$  galaxies.

#### 5. ELG selection with Fisher discriminant

We present in this section different target selections to account for different possible strategies. The eBOSS/ELG observations are planned to be done in the SGC and to begin in autumn 2016. A key point is the availability of the photometric data ahead of observations. The SDSS, SCUSS, and WISE data are already available, while the DECaLS data are in the process of acquisition and reduction. A first possibility is to make use of the maximum photometric information available today on the SGC, i.e. to combine the SDSS, SCUSS, and WISE data. A second possibility is to minimise the number of combined surveys, i.e. combining the SDSS and WISE data, to minimise possible systematics. A last possibility is to take advantage of the near availability of the DECam/ $z$  data, which will be two magnitudes deeper than the SDSS/ $z$  data. Indeed, the DECaLS/ $z$ -band imaging already covers more than  $1000 \text{ deg}^2$  over the SGC and should be made public through annual releases starting in March 2015, and the DES/ $z$ -band data over the “fat”-Stripe-82 region should be made public in 2015.

We first present in this section the tested ELG selection algorithms, which were built on the Fisher discriminant (Section 5.2). We present in Section 5.3 selections based on colour-colour cuts. The aim is two-fold: (1) We present the initially tested selection using  $ugri$ -bands to illustrate the improvement due to the addition of the WISE/W1 band; (2) the colour-colour cuts using WISE/W1 allow the comparison of the Fisher selections with classical methods. We then analyse the selection’s redshift and  $[OII]$  emission properties, along with their efficiencies (Section 5.4). This analysis is presented in two complementary manners to overcome the fact that the eboss6–7 test plates cover a biased subsample of some selections. In fact, the eboss6–7 test plates were designed to sample some selections, amongst which the Fisher\_UGRIZW1, Fisher\_GRIW1, and CC\_UGRI selections (defined in Sections 5.2 and 5.3). Those three selections thus have a  $\sim 95\%$  coverage with eboss6–7, since the remaining  $\sim 5\%$  were not targeted for tiling reasons and are thus an unbiased subsample. However, none of the ten eboss6–7 test plates have been specifically designed for the other tested selections: even if up to  $\sim 90\%$  of a selection is observed with eboss6–7, the unobserved  $\sim 10\%$  is a biased





**Fig. 4.**  $r - W1$  vs.  $g - r$  colour-colour plots illustrating how WISE helps in identifying  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs. *Left panel:* BC03 stellar population models; models with SFH  $\tau = 5, 10$  Gyr (SFH  $\tau = 0.05, 1.0$  Gyr, respectively) are representative of ELGs (LRGs, respectively); all models are dust-free ( $E(B - V) = 0$ ); and the reddening due to the dust ( $E(B - V) = 0.2$ ) is illustrated with the arrow in the top left corner. To guide the eye, we indicate with blue shading the approximate locus where  $0.6 \leq z_{\text{spec}} \leq 1.0$  ELGs are expected to lie according to those models. *Middle panel:* SDSS objects in the VVDS/F22 survey ( $17.5 \text{ mag} < i < 22.5 \text{ mag}$ ), binned in  $z_{\text{spec}}$ ; the hatched region is for stars. *Right panel:* SDSS  $0.6 \leq z_{\text{spec}} \leq 1.0$  galaxies in the VIPERS survey ( $i < 22.5 \text{ mag}$ ), binned in total [OII] luminosity. In the middle and right panels, we plot the region including 68% of the data for each bin.

subsample, thus preventing reliable statistics to be computed. For the redshift, we present results for eBOSS-7 spectroscopic redshifts (Fisher\_UGRIZW1, Fisher\_GRIW1, and CC\_UGRI selections) and for photometric redshifts (all selections). For the [OII] properties, we present results for the eBOSS-7 test plates (Fisher\_UGRIZW1, Fisher\_GRIW1, and CC\_UGRI) and for the VIPERS objects (all selections). Lastly, we present a brief illustration of the flexibility in terms of target density of our approach (Section 5.5).

### 5.1. Why use a Fisher discriminant approach?

The initially tested selection based on an SDSS-detection for eBOSS/ELG uses cuts in the  $uri$  and  $gri$  colour-colour diagrams. This selection, which we label CC\_UGRI (“CC” for colour-colour), is explained in Paper I and has been tested with the eBOSS6 observations. Our analysis deals with the five  $ugriz$  optical bands and the WISE/W1 near-infrared band, hence five independent colours. Though it is possible to define boxes in some colour-colour diagrams using those five colours, this task is complex and subjective, and this motivated us to use this alternative Fisher discriminant approach, which automatically defines the cut in the full colour-colour space. The only requirement is the availability of a training sample, which we have in hand thanks to the numerous spectroscopic data covering the SDSS footprint (see Section 3.2).

We thus present below the tested Fisher selections, and to compare their performance with classical colour-colour cuts, we also present two colour-colour cuts using the WISE/W1 band. In addition, we present the CC\_UGRI selection performance to illustrate the improvement due to the addition of the WISE/W1 band.

### 5.2. Fisher selection schemes

The strength of this Fisher selection scheme is its simplicity: our selections are solely based on the Fisher discriminant quantity ( $X_{FI, \min} < X_{FI}$ ) and on cuts in magnitudes and magnitude errors.  $X_{FI, \min}$  is tuned so that the selection has the desired object density. The tested selection schemes (including the Fisher discriminant definition) are described in Table 2.

As explained in Section 2.2, there are two possible approaches to defining the Signal and Background classes for the Fisher discriminant training. The first possible approach is to use criteria on both spectroscopic redshift  $z_{\text{spec}}$  and the total [OII] flux:

$$\begin{aligned} \text{Signal :} & \quad (0.6 < z_{\text{spec}} < 1.2 \text{ and } f_{\text{[OII]}}^{\text{tot}} > 10^{-16.1} \text{ erg.s}^{-1} \text{ cm}^{-2}) \\ & \quad \text{or } (0.9 < z_{\text{spec}} < 1.2 \text{ and } g < 22 \text{ mag}) \\ \text{Background :} & \quad z_{\text{spec}} < 0.5. \end{aligned} \quad (3)$$

With this, we account for the possibility that in our training sample, we miss some [OII] line measurement due to sky lines at  $z_{\text{spec}} > 0.9$ , while keeping objects at  $0.9 < z_{\text{spec}} < 1.2$  in our Signal class with significant flux in the  $g$  band, since it is correlated with the [OII] flux (see Comparat et al. 2015b). Because we are working with objects detected in the SDSS, the vast majority of the galaxies we want to exclude will be at low redshift, so we only include those in our Background class (regardless of the [OII] flux). The second possible approach is to define our Signal class using only the  $z_{\text{spec}}$  quantity:

$$\begin{aligned} \text{Signal :} & \quad 0.75 < z_{\text{spec}} < 1.3 \\ \text{Background :} & \quad z_{\text{spec}} < 0.5. \end{aligned} \quad (4)$$

We tested five selections, based on three different survey combinations:

- Fisher\_UGRIZW1: SCUSS/ $u$  + SDSS/ $griz$  + WISE/W1;
- Fisher\_GRIW1: SDSS/ $gri$  + WISE/W1;
- Fisher\_GRIW1<sup>OII</sup>: SDSS/ $gri$  + WISE/W1, with a Fisher training using Eq.(3);
- Fisher\_GRZW1<sub>180</sub>: SDSS/ $gr$  + CFHTLS-W/ $z$  (degraded to the DECaLS depth) + WISE/W1;
- Fisher\_GRZW1<sub>300</sub>: same as Fisher\_GRZW1<sub>180</sub>, but with higher object density.

Except for the Fisher\_GRIW1<sup>OII</sup> selection, all the Fisher trainings were done with Eq.(4). Our choice to define the tested selections were guided by the eBOSS/ELG experiment requirements (number of targets, available imaging at the start of spectroscopic observations). On the one hand, the Fisher\_UGRIZW1 selection is based on a broad wavelength coverage (from the  $u$  band to the  $W$ ) and is limited to the SGC: as mentioned in

	Fisher_UGRIZW1	Fisher_GRIW1	Fisher_GRIW1 <sup>OH</sup>	Fisher_GRZW1 <sub>180</sub> , Fisher_GRZW1 <sub>300</sub>
Photometry	SCUSS/ <i>u</i> SDSS/ <i>griz</i> WISE/ <i>W1</i>		SDSS/ <i>gri</i> WISE/ <i>W1</i>	SDSS/ <i>gr</i> DECaLS-like/ <i>z</i> WISE/ <i>W1</i>
$X_{FI}$ training	$z_{\text{spec}}$ only (Eq. 4)	$z_{\text{spec}}$ only (Eq. 4)	$z_{\text{spec}}$ and $f_{[\text{OII}]}$ (Eq. 3)	$z_{\text{spec}}$ only (Eq. 4)
$X_{FI}$ definition	$X_{FI} = \alpha_0 + \alpha_{ur} \times (u - r) + \alpha_{gr} \times (g - r) + \alpha_{ri} \times (r - i) + \alpha_{rz} \times (r - z) + \alpha_{rW1} \times (r - W1)$			
	$\begin{cases} \alpha_0 = +0.956 \\ \alpha_{ur} = -0.650 \\ \alpha_{gr} = -0.781 \\ \alpha_{ri} = +0.065 \\ \alpha_{rz} = +0.229 \\ \alpha_{rW1} = +0.739 \end{cases}$	$\begin{cases} \alpha_0 = +0.104 \\ \alpha_{ur} = 0 \\ \alpha_{gr} = -1.308 \\ \alpha_{ri} = +0.870 \\ \alpha_{rz} = 0 \\ \alpha_{rW1} = +0.782 \end{cases}$	$\begin{cases} \alpha_0 = +1.103 \\ \alpha_{ur} = 0 \\ \alpha_{gr} = -1.982 \\ \alpha_{ri} = +0.298 \\ \alpha_{rz} = 0 \\ \alpha_{rW1} = +0.701 \end{cases}$	$\begin{cases} \alpha_0 = +0.519 \\ \alpha_{ur} = 0 \\ \alpha_{gr} = -1.483 \\ \alpha_{ri} = 0 \\ \alpha_{rz} = -0.120 \\ \alpha_{rW1} = +0.967 \end{cases}$
$X_{FI}$ cuts	$1.321 < X_{FI}$	$1.492 < X_{FI}$	$1.544 < X_{FI}$	$1.475, 1.141^\dagger < X_{FI}$
Magnitude cuts	$\begin{cases} 20.0 < u < 23.5, u_{\text{err}} < 1.0 \\ 20.0 < g < 22.5, g_{\text{err}} < 0.5 \\ 19.0 < r < 22.5, r_{\text{err}} < 0.5 \\ 19.0 < i < 21.5, i_{\text{err}} < 0.5 \\ 17.0 < W1 < 21.0, W1_{\text{err}} < 0.5 \end{cases}$	$\begin{cases} 20.0 < g < 22.5, g_{\text{err}} < 0.5 \\ 19.0 < r < 22.5, r_{\text{err}} < 0.5 \\ 19.0 < i < 21.5, i_{\text{err}} < 0.5 \\ 17.0 < W1 < 21.0, W1_{\text{err}} < 0.5 \end{cases}$		$\begin{cases} 20.0 < g < 22.5, g_{\text{err}} < 0.5 \\ 19.0 < r < 22.5, r_{\text{err}} < 0.5 \\ 18.0 < z < 21.5, z_{\text{err}} < 0.5 \\ 17.0 < W1 < 21.0, W1_{\text{err}} < 0.5 \end{cases}$
Other cuts (for all selections)		BINNED2 = 0 OBJC.TYPE = 3 or $r > 22$ SDSS photometric masks ( <code>bright_star</code> , <code>bad_field</code> , <code>bright_object_rykoff</code> ) and custom <i>W1</i> mask		

**Table 2.** Criteria used to define our Fisher selections. For the Fisher\_GRZW1 selection, the  $^\dagger$  denotes quantities related to the deep (300 objects per  $\text{deg}^2$ ) selection.

Section 3.1, the SCUSS deep *u*-band photometry provides us with a measurement of the ultra-violet emission at redshifts 0.6–1.0. On the other hand, the Fisher\_GRIW1 selection has a narrower wavelength coverage, but has the advantages of being available on both the SGC and the NGC and of minimising the number of surveys used. Eventually, the Fisher\_GRZW1 selections use a deeper – hence less scattered – *z*-band photometry. We tested two scenarios (target densities of  $180 \text{ deg}^{-2}$  and  $300 \text{ deg}^{-2}$ ) to see to what extent the target density can be increased, in case the available DECaLS *z*-band photometry is available for an area smaller than  $1500 \text{ deg}^2$ .

Figure 5 illustrates how the Fisher correlates with redshift and [OII] flux. The VIPERS sample, with [OII] flux measurement in  $0.6 \lesssim z_{\text{spec}} \lesssim 1.2$ , allows us to see the simultaneous dependence on  $z_{\text{spec}}$  and  $F_{[\text{OII}]}^{\text{tot}}$ , whereas we use the VVDS/F22 sample to probe the dependence only on  $z_{\text{spec}}$ , but over a wide redshift range ( $0 < z_{\text{spec}} < 1.2$ ). We recall that the data plotted here are the subsamples of the VIPERS and VVDS/F22 detected in the SDSS and that, due to the SDSS depth, the number of galaxies per redshift bin decreases for  $z_{\text{spec}} \gtrsim 0.6$ . The Fisher\_UGRIZW1 quantity correlates with both redshift and [OII] flux, which means that selecting objects with high values of Fisher\_UGRIZW1 should be efficient in selecting ELGs in  $0.6 \leq z \leq 1.0$ . We note that the Fisher training here is done on  $z_{\text{spec}}$  only (Eq. (4)): the efficiency in selecting [OII] emitters is a byproduct of the presence of the deep SCUSS *u* band. Indeed, the faint magnitude cut on the SCUSS photometry favours  $z \sim 0.8$  star-forming galaxies against  $z \sim 0.8$  passive galaxies, which have faint emission in the *u* band (see for instance Comparat et al. 2015b). In contrast, we observe that, though having a strong correlation with redshift by training, the Fisher\_GRIW1 is inefficient at distinguishing [OII] emitters. Adding [OII] in the training (Fisher\_GRIW1<sup>OH</sup>, trained with Eq.(3)) improves the efficiency in selecting [OII] emitters, but at the cost of slightly reducing the correlation with redshift.

Lastly, the Fisher\_GRZW1 efficiently distinguishes the redshift, but favours low [OII] emitters at  $z_{\text{spec}} \sim 1$ .

Before applying the cut on the Fisher discriminant, we apply cuts on the magnitudes and their errors (see Table 2). The aim is twofold: the cut on the bright magnitudes removes objects that surely are at low redshift ( $z \leq 0.6$ ) from the samples. The cut on the faint magnitudes and on the magnitude errors ensures that we use reasonable photometry and mitigates inhomogeneous spatial distribution. We note that the Fisher discriminant’s dependence on redshift and [OII] flux illustrated in Figure 5 does not change when those cuts are applied, though the sample is significantly smaller.

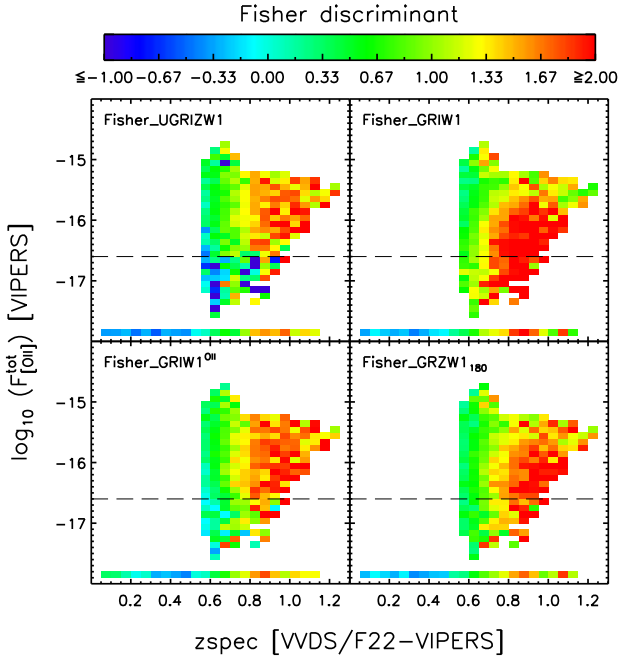
We also apply the following cuts: 1) we reject potential stars, 2) we reject objects lying in regions where the photometry is uncertain, 3) we reject objects with  $\text{BINNED2} \neq 0$ . To reject potential stars, we reject objects with  $\text{OBJC.TYPE} \neq 3$  and  $r < 22$ , where  $\text{OBJC.TYPE}$  is the SDSS star/galaxy separator<sup>8</sup>). Regarding the regions where the photometry is uncertain, we use the SDSS masks (`bright_star`, `bad_field`, `bright_object_rykoff`; see <http://data.sdss3.org/sas/dr10/boos/lss/>) and a custom mask for the *W1* bright-object neighbourhood. In fact, as explained in Lang et al. (2014, see their Figure 10), SDSS objects falling in *W1* halo outskirts are not masked and have a significantly overestimated *W1* flux – so are very red in  $r - W1$ .

To build our custom *W1* mask, we detect spatial  $5\sigma$  overdensities of SDSS  $r < 22.5$  mag objects with  $3 \text{ mag} < r - W1 < 5 \text{ mag}$  and mask objects within a radius of  $0.4^\circ$  around those overdensities<sup>9</sup>. Our custom *W1* mask includes  $\sim 200$  overden-

<sup>8</sup> [http://www.sdss.org/dr12/algorithms/classify/#photo\\_class](http://www.sdss.org/dr12/algorithms/classify/#photo_class)

<sup>9</sup> To detect overdensities, we first create a density map with a  $0.1^\circ \times 0.1^\circ$  binning and then run SExtractor (Bertin & Arnouts 1996) with  $\text{DETECT\_MINAREA} = 10$  and  $\text{DETECT\_THRESH} = 5$ .





**Fig. 5.** Dependence of the four defined Fisher discriminants on  $z_{\text{spec}}$  and  $F_{\text{[OIII]}}^{\text{tot}}$ . For each Fisher discriminant, we present the VIPERS data (main patch), while the VWDS/F22 data are displayed at  $\log_{10}(F_{\text{[OIII]}}^{\text{tot}}) = -17.8$  in  $z_{\text{spec}}$  bin only to illustrate the redshift dependence over  $0 < z_{\text{spec}} < 1.2$ . The horizontal dashed line represents the approximate  $F_{\text{[OIII]}}^{\text{tot}}$  of objects with  $S/N = 3$ . We recall that we only consider galaxies detected in the SDSS. We only display bins with at least ten galaxies.

sities in the SGC (and  $\sim 350$  in the NGC). We note that the masking is independent of any ELG selection. We also notice that, though the CFHTLS/W1 region used in this study does not include bright W1 objects, such a W1 mask is nevertheless required when using our selections over larger regions, as for instance in Paper III, where the homogeneity and systematics analysis of our selections over the SGC is done. Eventually we only keep objects with  $\text{BINNED2} = 0$ , because this (slightly) reduces the number of objects with unexploitable spectra. ( $\text{BINNED2}$  is one of the SDSS photometric flags<sup>10</sup>.)

### 5.3. Colour-colour cut selections

We also present some selections using colour-colour cuts in the following analysis. First, we present the CC\_UGRI selection, which uses cuts in the  $uri$  and  $gri$  colour-colour diagrams. This selection is detailed in Paper I and has been tested with the eboss6 observations. It allows the improvement in the selection to be quantified thanks to the addition of the WISE/W1 band.

Additionally, we present two selections based on colour-colour cuts using the  $ugrizW1$  bandset and the  $griW1$  bandset, which we label CC\_UGRIZW1 and CC\_GRIW1. Those allow a comparison of the Fisher discriminant selection performances with classical colour-colour cuts. The CC\_UGRIZW1 selection

has cuts similar to the Fisher\_UGRIZW1 (see Table 2) selection except the cut on the Fisher discriminant, which is replaced by

$$\begin{aligned} g-r &> 2.0 \times (u-r) - 3.00, \\ g-r &< 1.2 \times (r-i), \\ i-z &> -2.4 \times (r-i) + 0.60, \text{ and} \\ r-W1 &> 2.0 \times (g-r) + 0.35. \end{aligned} \quad (5)$$

The CC\_GRIW1 selection has cuts similar to the Fisher\_GRIW1 (see Table 2) selection except the cut on the Fisher discriminant, which is replaced by

$$\begin{aligned} g-r &< 1.0, \\ r-i &> 0.5, \text{ and} \\ r-W1 &> 2.5 \times (g-r) + 0.25. \end{aligned} \quad (6)$$

### 5.4. Selection properties

We recall that the redshift requirements for a  $180 \text{ deg}^{-2}$  eBOSS/ELG target selection are (1) an efficiency of 70% in the  $0.6 \leq z \leq 1.0$  range (i.e. at least 70% of the targets have a measurable  $z_{\text{spec}}$  with  $0.6 \leq z \leq 1.0$ ); (2) a redshift failure rate  $\lesssim 1\%$  in  $0.6 \leq z \leq 1.0$ .

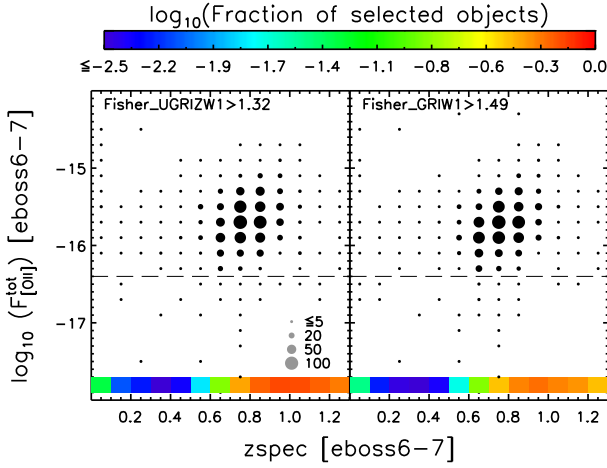
Figure 6 represents the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections in the [OIII] flux versus redshift diagram using our eboss6-7 test plates measurements. For each redshift bin (using  $z_{\text{phot}}$ ), we colour-code at  $\log_{10}(F_{\text{[OIII]}}^{\text{tot}}) = -17.8$  the ratio of the number of selected photometric objects to the number of photometric objects passing all the cuts in Table 2, except the cut on the Fisher discriminant. We do not display the Fisher\_GRIW1<sup>OII</sup> and Fisher\_GRZW1 selections because their eboss6-7 observations are biased. As expected from the preliminary analysis of Figure 5, the two Fisher selections are efficient in selecting  $0.6 \leq z \leq 1.0$  galaxies with a significant [OIII] flux. More precisely, we see that the cut on the Fisher discriminant is very efficient in rejecting  $z_{\text{spec}} < 0.6$  objects, as it rejects  $> 99\%$  of those.

We now study the magnitude and colours distributions for our tested selections, then the redshift distributions and the [OIII] emission; we finally present the overall statistics for all the tested selections.

#### 5.4.1. Magnitudes and colours

Figure 7 summarises the magnitude and colour distributions of the selections, along with the percentage of selected objects with  $0.6 \leq z_{\text{phot}} \leq 1.0$ . Overall, the five Fisher selections have broadly similar magnitudes distributions, except for the  $u$  band and the DECaLS/ $z$ -band, where we apply different magnitude cuts. When compared to the four other Fisher selections, the Fisher\_UGRIZW1 selection almost has no objects with red  $u-r > 2$  colours, which are likely to be more passive galaxies. One reason for that is the presence of an upper  $u$ -band magnitude limit in the selection criteria, which requires a minimum flux in the  $u$  band. For all colours, the Fisher\_GRIW1<sup>OII</sup> selection has a distribution that is bluer than the Fisher\_GRIW1 selection one, confirming the link between blue colours and [OIII] emission. When compared to the Fisher\_GRZW1<sub>180</sub> selection, the Fisher\_GRZW1<sub>300</sub> selection shows small differences; for instance, it has slightly redder  $g-r$  colours and slightly bluer  $r-W1$  colours, consistent with a lower redshift (see Figure 4). Additionally, all our Fisher selections show common trends in the percentage of selected objects with  $0.6 \leq z_{\text{phot}} \leq 1.0$ . For instance, the selected objects with faint  $u$  or  $g$  magnitudes are,

<sup>10</sup> <http://skyserver.sdss.org/dr9/en/help/browser/enum.asp?n=PhotoFlags> for the large majority, in our desired redshift range; however,



**Fig. 6.** Dependence of the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections on  $z_{\text{spec}}$  and  $F_{\text{[OII]}}^{\text{tot}}$  for our observed eboss6-7 plates. The dot size scales with the number of objects entering the bin. For each redshift bin (using  $z_{\text{phot}}$ ), we colour-coded at  $\log_{10}(F_{\text{[OII]}}^{\text{tot}}) = -17.8$  the ratio of the number of selected photometric objects to the number of photometric objects passing all the cuts in Table 2 except the cut on the Fisher discriminant. The horizontal dashed line represents the approximate  $F_{\text{[OII]}}^{\text{tot}}$  of objects with  $S/N = 3$ .

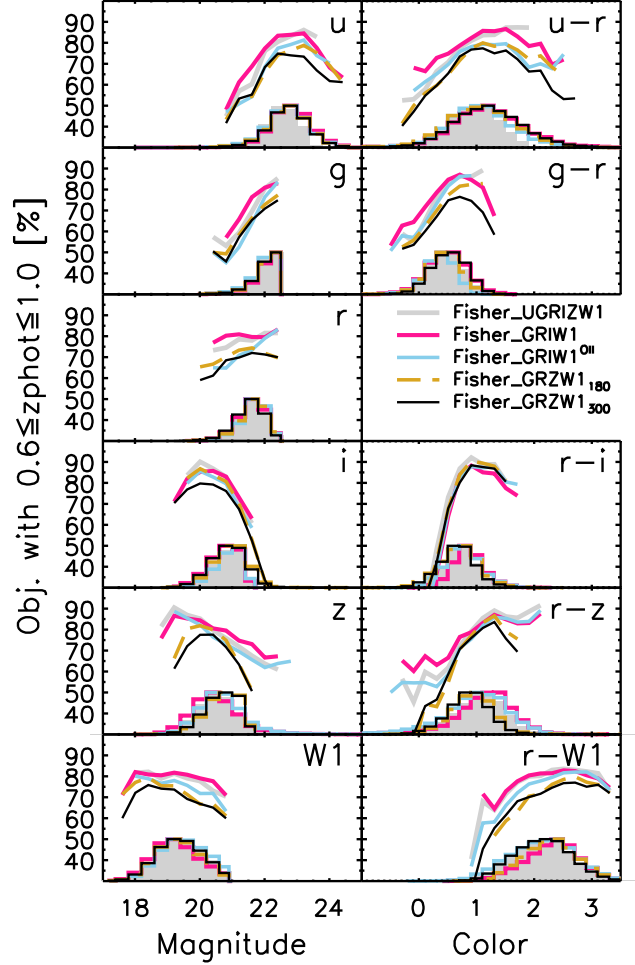
the depth of the images and the requirement to obtain an usable spectrum with a 1h observation prevent us from pushing the selection at fainter magnitudes.

In Section 4, we studied the location of the  $0.6 \leq z_{\text{phot}} \leq 1.0$  ELGs in the  $r - W1$  vs.  $g - r$  colour-colour diagram. We look in Figure 8 at the same diagram for our tested Fisher selections: overall, the selected galaxies are indeed located in the expected region. We can see some small differences in the loci occupied by the different selections. For instance, the approximate cut in this diagram has a steeper slope for the Fisher\_UGRIZW1 and Fisher\_GRIW1<sup>OH</sup> selections than for the Fisher\_GRIW1 and Fisher\_GRZW1<sub>180</sub> selections, and this steeper slope implies the selection of more galaxies at  $(g - r, r - W1) \sim (0, 1)$  that have higher redshift and are strongly star-forming, while the flatter slope implies the selection of more galaxies at  $(g - r, r - W1) \sim (1, 2)$ , with less star formation.

#### 5.4.2. Redshift

We now study the redshift distributions for the Fisher and colour-colour selections. The top panel of the Figure 9 represents the  $z_{\text{spec}}$  distribution from eboss6-7 observations for the three selections that have unbiased sampling. The middle (bottom, respectively) panel represents the CFHTLS/ $z_{\text{phot}}$  distribution for the Fisher selections (colour-colour selections, respectively). We observe that the eboss6-7  $z_{\text{spec}}$  distributions for the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections are close, i.e. peaked between  $0.6 \leq z_{\text{phot}} \leq 1.0$  with little contamination from  $z_{\text{spec}} < 0.6$  objects. The Fisher\_UGRIZW1 selection has a slightly higher median  $z_{\text{spec}}$  than the Fisher\_GRIW1 selection. The CC\_UGRI selection has a distribution that is slightly shifted to lower redshifts with more contamination from  $z_{\text{spec}} < 0.6$  objects.

We verify in the middle and bottom panels that the  $z_{\text{phot}}$  distributions faithfully mimic the  $z_{\text{spec}}$  distributions for the three



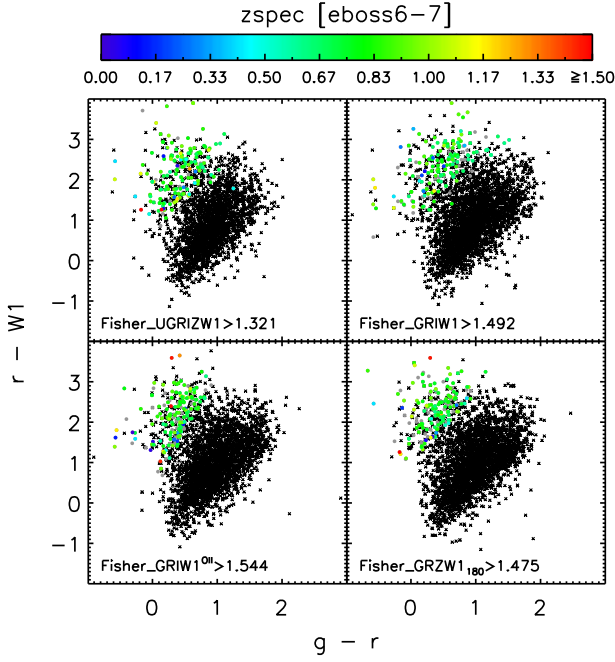
**Fig. 7.** Photometric properties of the five Fisher and the CC\_UGRI selections. For each magnitude and colour, we report the percentage of selected objects having  $0.6 \leq z_{\text{phot}} \leq 1.0$ , along with the magnitude normalised distribution of the selection. For the Fisher\_GRZW1 selections, the  $z$  band corresponds to the DECaLS-like photometry obtained from the CFHTLS-Wide  $z$ -band photometry.

selections of the top panel, though with a slight shift towards higher values of redshift at  $\sim 1$ , as expected from Appendix A. Furthermore, we see that the Fisher\_GRIW1<sup>OH</sup> selection has more contamination from  $z_{\text{phot}} < 0.6$  objects than the Fisher\_GRIW1 selection. The Fisher\_GRZW1<sub>180</sub> selection has slightly more galaxies at  $z > 1$ , while the Fisher\_GRZW1<sub>300</sub> selection has a distribution peaking at a lower redshift and has more  $z_{\text{phot}} < 0.6$  contamination than the Fisher\_GRZW1<sub>180</sub> selection.

The colour-colour selections including the W1 band have  $z_{\text{phot}}$  distributions comparable to the Fisher selections. We see that including the W1 band reduces the contamination from  $z_{\text{phot}} < 0.6$  objects.

#### 5.4.3. [OII] flux

In Figure 10 we display the total [OII] flux distributions, and in Figure 11 the percentage of objects with  $f_{\text{[OII]}}^{\text{tot}} \geq 10^{-16}$



**Fig. 8.**  $r - W1$  vs.  $g - r$  colour-colour plot of the four defined Fisher discriminants. For each selection, crosses represent photometric objects in the eboss6-7 area passing all the cuts in Table 2, except the cut on the Fisher discriminant. Filled dots represent those passing the cut on the Fisher discriminant with the colour coding  $z_{\text{spec}}$  when observed with the eboss6-7 test plates (grey colour when no reliable  $z_{\text{spec}}$  could be estimated from the spectrum or if the object has not been observed). For clarity, we display only one out of ten objects.

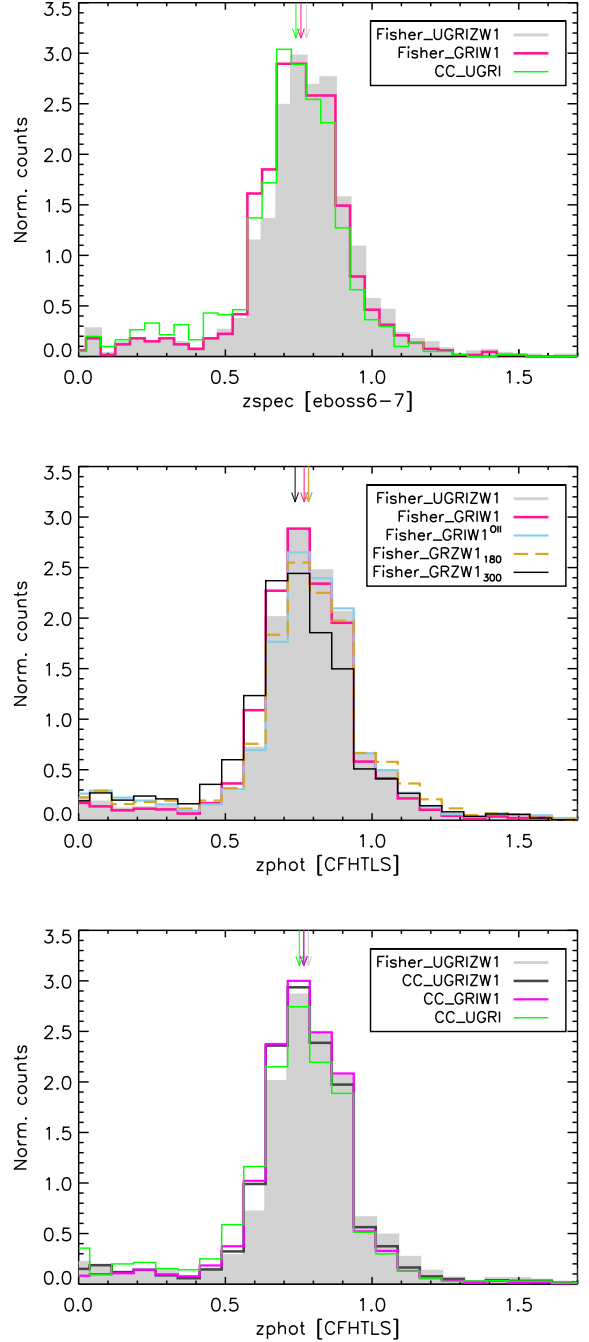
$\text{erg.cm}^{-2}.\text{s}^{-1}.\text{\AA}^{-1}$  as a function of the  $g$ -band magnitude. Total fluxes above  $10^{-16} \text{ erg.cm}^{-2}.\text{s}^{-1}.\text{\AA}^{-1}$  are detected at more than  $5\sigma$  in both the VIPERS and eBOSS. This threshold allows us to directly compare the percentages using the VIPERS and the eBOSS, overcoming the fact the VIPERS spectra have a slightly higher signal-to-noise ratio than the eBOSS spectra.

The Fisher\_UGRIZW1 selection emits more [OII] than does the Fisher\_GRIW1 selection. This is consistently seen from eboss6-7 and VIPERS data. When looking at the VIPERS data, we see that, as expected from the Fisher training, the Fisher\_GRIW1<sup>0II</sup> selection is more [OII]-emitting than the Fisher\_GRIW1 selection. The Fisher\_GRZW1 selection is intermediate, and the [OII] emission is similar for both tested densities (180 and 300  $\text{deg}^{-2}$ ). The VIPERS data also show that the colour-colour selection has slightly less [OII] emission than the Fisher\_UGRIZW1 selection. Noticeably, the CC\_UGRI selection is significantly less [OII]-emitting than the other selections.

Finally, we see the trend for all selections and eboss6-7 and VIPERS data towards galaxies that are fainter in the  $g$  band to have less [OII] emission on average. This result agrees with those from the Comparat et al. (2015b) study of the [OII] luminosity function.

#### 5.4.4. Statistics and summary

We give details on the properties of our tested Fisher selections in Table 3 and of the colour-colour selections in Table 4.



**Fig. 9.** Redshift distributions for the five Fisher and the CC\_UGRI selections. The median redshift value is indicated by an arrow on the top x-axis. *Top panel:* eboss6-7  $z_{\text{spec}}$ . *Middle panel:* CFHTLS  $z_{\text{phot}}$  for the Fisher selections. *Bottom panel:* CFHTLS  $z_{\text{phot}}$  for the colour-colour selections. We report the Fisher\_UGRIZW1 selection to facilitate the comparison.

We present the statistics for redshift, [OII] flux, and overall efficiency. More precisely, lines (L1-L6) present information computed with the photometric data over a  $\sim 50 \text{ deg}^2$  area within the CFHTLS/W1 field (density, overlap with LRGs, photometric redshifts statistics); lines (L7-L8) report the numbers of  $0.6 < z_{\text{spec}} < 1.0$  VIPERS galaxies passing the selection, along with the percentage of those having  $f_{\text{[OII]}}^{\text{tot}} > 10^{-16}$



erg.cm<sup>-1</sup>.s<sup>-2</sup>. Lines (L9-L14) present spectroscopic information for the eBOSS6-7 plates observations covering an area of 8.82 deg<sup>2</sup> (number of galaxies, percentage of observed galaxies, percentage of galaxies with non reliable  $z_{\text{spec}}$  measurement,  $z_{\text{spec}}$  statistics). Finally, lines (L15-L16) present spectroscopic information for the eBOSS6-7 galaxies with a reliable  $z_{\text{spec}}$  measurement with  $0.6 \leq z \leq 1.0$ , which are those objects that would be used for a BAO measurement (mean  $z_{\text{spec}}$ , percentage of galaxies with  $f_{\text{[OII]}}^{\text{tot}} > 10^{-16}$  erg.cm<sup>-1</sup>.s<sup>-2</sup>, expected percentage of galaxies with catastrophic  $z_{\text{spec}}$  estimation). We now summarise the selection properties based on the above analysis of Figures 6-11 and statistics from Tables 3-4.

**Fisher\_UGRIZW1 selection.** This selection meets the initial eBOSS/ELG target selection redshift criteria. It has an efficiency of 71% and an expected  $z_{\text{spec}}$  failure rate of 0.5%. We observe that it has a narrow  $z_{\text{spec}}$  distribution with a typical width of 0.12. The  $z_{\text{phot}}$  distribution is very similar. The median  $z_{\text{spec}}$  of the selection is 0.78. As expected from the preliminary study (Figure 5), the Fisher\_UGRIZW1 selection is efficient in selecting [OII] emitters, and this can be seen in particular in Figure 11 and also in the lower value of objects with an unreliable  $z_{\text{spec}}$  or with the small overlap with LRGs (12% and 2.0 deg<sup>-2</sup>, respectively; see Table 3).

**Fisher\_GRIW1 and Fisher\_GRIW1<sup>OII</sup> selections.** The Fisher\_GRIW1 selection also meets the initial eBOSS/ELG target selection redshift criteria, the efficiency and expected  $z_{\text{spec}}$  failure rate in  $0.6 \leq z_{\text{spec}} \leq 1.0$  being close to the requirements (71% and 0.9%, respectively). The shape of the redshift distribution is close to the one of the Fisher\_UGRIZW1 selection, but shifted to a slightly lower value (0.76 vs. 0.78 for  $z_{\text{spec}}$ ). The Fisher\_GRIW1 selection is also a little more efficient at removing low-redshift objects, as expected from Figure 5, where we see a strong correlation between  $z_{\text{spec}}$  and the Fisher discriminant. These features are also visible in the  $z_{\text{phot}}$  distribution. An important characteristic of the Fisher\_GRIW1 selection is that it tends to select fewer [OII] emitters than the Fisher\_UGRIZW1 selection. Figure 11 consistently supports this observation, using both the eBOSS and the VIPERS measurements; in addition, this can also be seen in Table 3, where the number of LRGs per square degree is almost twice higher than for the Fisher\_UGRIZW1 selection, or in Figure 7, where the selected objects have redder  $u - r$  colours. The Fisher\_GRIW1<sup>OII</sup> selection succeeds to select more [OII] emitters (middle panels of Figures 10 and 11, and overlap with LRGs in Table 3), but at the cost of being less efficient at removing low-redshift objects (middle panel of Figure 9). For example, it can be seen in Figures 7 and 8 that the Fisher\_GRIW1<sup>OII</sup> selection has colours bluer than in the Fisher\_GRIW1 selection. Unfortunately, even if we probe 89% of the Fisher\_GRIW1<sup>OII</sup> selection with the eBOSS6-7 test plates, we cannot infer robust statistics with them because ~5% of the untargeted objects are a biased subsample. Typically, low-redshift objects belonging to the Fisher\_GRIW1<sup>OII</sup> selection but not to the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections will not be targeted.

**Fisher\_GRZW1<sub>180</sub> and Fisher\_GRZW1<sub>300</sub> selections.** The Fisher\_GRZW1<sub>180</sub> selection offers an interesting alternative. The photometric redshift distribution is similar to the one of the Fisher\_UGRIZW1 selection, though it includes a hardly higher number of  $z_{\text{phot}} \leq 0.2$  galaxies (see Table 3 and middle panel of Figure 9). In terms of [OII] emission, the VIPERS data indicate

that this selection lies midway between the Fisher\_UGRIZW1 and the Fisher\_GRIW1 selections. If we increase the target density to 300 deg<sup>-2</sup>, the  $z_{\text{phot}}$  distribution broadens and is shifted to lower values (0.74 vs. 0.78), but the Fisher\_GRZW1<sub>300</sub> selection has similar [OII] emission properties to a target density of 180 deg<sup>-2</sup>.

**Colour-colour selections.** With the eBOSS6-7 data, we see that the CC\_UGRI selection has a lower median redshift (0.74) and low efficiency (59%), because this selection emits less [OII], implying more galaxies without exploitable redshift because of the lack of significant emission lines. Thus, the WISE/W1 data help in increasing the selection redshift and [OII] flux, hence in increasing the selection efficiency. The other two tested selections (CC\_UGRIZW1 and CC\_GRIW1) have characteristics that are broadly similar to the corresponding Fisher selections. More precisely, on the one hand, they have slightly more objects with  $0.6 \leq z_{\text{phot}} \leq 1.0$ , but on the other, their median redshift is slightly lower, and they have slightly less [OII] emission.

### 5.5. Adjusting the selection density

Our tests in Section 5.4 were done on a single location in a rather small area compared to the aim of 1500 deg<sup>2</sup> of the eBOSS/ELG survey. Here we investigate 1) the mean object density over large SGC areas and 2) the way the selection efficiency varies if we change the selection density. We refer to Paper III for a complete analysis over the full SGC.

The lower cut on the Fisher discriminant,  $X_{FI,\text{min}}$ , is set so that our selections have an object density of 180 deg<sup>-2</sup> over a ~50 deg<sup>2</sup> area included in the CFHTLS-Wide W1 field, approximately centred at R.A.=34 and DEC.=-6.5. Thanks to the recent development of catalogue tools for the Paper III analysis, it is now feasible to apply our selections using SDSS, WISE, and SCUSS photometry over larger SDSS footprints. We computed the object density for two typical SGC areas of ~700 deg<sup>2</sup> each: one that we label LowDec ( $-35 < \text{R.A.} < 40$  and  $-5 < \text{DEC.} < 5$ ) and one labelled HiDec ( $0 < \text{R.A.} < 30$  and  $5 < \text{DEC.} < 30$ ).

The Fisher\_UGRIZW1 selection has a mean object density of 183 deg<sup>-2</sup> (166 deg<sup>-2</sup>, respectively) over the LowDec (HiDec, respectively) area, whereas the Fisher\_GRIW1 and the Fisher\_GRIW1<sup>OII</sup> selections have a mean object density of 183-187 deg<sup>-2</sup> over both areas. The Fisher\_UGRIZW1 thus seems to have a less homogeneous density over the SGC than the Fisher\_GRIW1 and Fisher\_GRIW1<sup>OII</sup> selections.

We note that the Fisher\_UGRIZW1 selection object density can be increased by lowering  $X_{FI,\text{min}}$ , the threshold cut on the Fisher discriminant. To illustrate this flexibility in our method, we do the following exercise using the CFHTLS  $z_{\text{phot}}$  for our ~50 deg<sup>2</sup> test area. We look at the variations of the mean  $z_{\text{phot}}$  and of the percentage of galaxies with  $0.6 \leq z_{\text{phot}} \leq 1.0$  as a function of the selection density, when varying the Fisher discriminant threshold cut  $X_{FI,\text{min}}$  from 2.0 to 1.0 in steps of 0.1. We note that the efficiencies computed with the  $z_{\text{phot}}$  are higher than the ones computed with the  $z_{\text{spec}}$  because there is no requirement to have a measurable redshift from the spectrum; however, we are here interested in the relative variation with the threshold cut on the Fisher discriminant. The results are displayed in Figure 12, where we see that the four Fisher discriminants have similar be-

	Fisher_UGRIZW1	Fisher_GRIW1	Fisher_GRIW1 <sup>OH</sup>	Fisher_GRZW1 <sub>180</sub>	Fisher_GRZW1 <sub>300</sub>
<b>Photometric data (~50 deg<sup>2</sup>)</b>					
(L1) Density (deg <sup>-2</sup> )	180	182	181	183	301
(L2) eBOSS/LRG overlap (deg <sup>-2</sup> )	2.0	3.5	2.4	2.7	3.4
(L3) median( $z_{\text{phot}}$ )	0.78	0.77	0.78	0.78	0.74
(L4) $z_{\text{phot}}$ peak width <sup>†</sup>	0.12	0.13	0.12	0.13	0.14
(L5) % with $0.6 \leq z_{\text{phot}} \leq 1.0$	79%	80%	75%	73%	71%
(L6) mean( $z_{\text{phot}}$ ) ( $0.6 \leq z_{\text{phot}} \leq 1.0$ )	0.79	0.78	0.79	0.79	0.77
<b>VIPERS (<math>0.6 \leq z_{\text{spec}} \leq 1.0</math>)</b>					
(L7) N selected galaxies	555	552	512	513	874
(L8) $f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$	78%	69%	76%	74%	75%
<b>Plates eboss6-7 (8.82 deg<sup>2</sup>)</b>					
(L9) N selected galaxies	1609	1685	1684	1669	2763
(L10) Targeted	92%	93%	89%	83%	68%
(L11) Unreliable $z_{\text{spec}}$	12%	15%	(15%)	(13%)	(13%)
(L12) Median( $z_{\text{spec}}^{\text{reliable}}$ )	0.78	0.76	(0.78)	(0.78)	(0.76)
(L13) $z_{\text{spec}}^{\text{reliable}}$ peak width <sup>†</sup>	0.12	0.12	(0.12)	(0.12)	(0.13)
(L14) Efficiency ( $0.6 < z_{\text{spec}}^{\text{reliable}} < 1.0$ )	71%	71%	(67%)	(71%)	(71%)
<b>Plates eboss6-7, <math>0.6 &lt; z_{\text{spec}}^{\text{reliable}} &lt; 1.0</math> only</b>					
(L15) Mean( $z_{\text{spec}}^{\text{reliable}}$ )	0.79	0.77	(0.79)	(0.78)	(0.77)
(L16) $f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$	85%	81%	(84%)	(81%)	(82%)
(L17) Expected $z_{\text{spec}}$ failure	0.5%	0.9%	(0.7%)	(0.8%)	(0.8%)

**Table 3.** Summary of the Fisher selection properties. Lines (L1-L6): information computed with the photometric data over a ~50 deg<sup>2</sup> area within the CFHTLS/W1 field (density, overlap with LRGs identified using the cuts defined in Prakash et al. (2015), photometric redshifts statistics). Lines (L7-L8): number of  $0.6 \leq z_{\text{spec}} \leq 1.0$  VIPERS galaxies passing the selection, and percentage of those having  $f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$ . Lines (L9-L14): spectroscopic information for the eboss6-7 plates observations covering an area of 8.82 deg<sup>2</sup> (number of galaxies, percentage of observed galaxies, percentage of galaxies with unreliable  $z_{\text{spec}}$  measurement,  $z_{\text{spec}}$  statistics). Lines (L15-L16) present spectroscopic information for the eboss6-7 galaxies with a reliable  $z_{\text{spec}}$  measurement with  $0.6 \leq z \leq 1.0$  (mean  $z_{\text{spec}}$ , percentage of galaxies with  $f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$ , expected percentage of galaxies with catastrophic  $z_{\text{spec}}$  estimation). For lines (L10-L16), we report in brackets the quantities derived from our spectroscopic observations for the Fisher\_GRIW1<sup>OH</sup> and Fisher\_GRZW1 selections: those quantities are biased because they are obtained from a non-random subsample constituted of objects passing the Fisher\_UGRIZW1 or Fisher\_GRIW1 selections. <sup>†</sup>: the width is estimated through the fitting of Gaussian.

haviours. If we decrease the threshold cut on the Fisher discriminant (from left to right), the density increases and the mean  $z_{\text{phot}}$  decreases, meaning that we select more galaxies but they have a lower redshift. The percentage of galaxies with  $0.6 \leq z_{\text{phot}} \leq 1.0$  decreases when the density increases, because there are more selected galaxies at  $z_{\text{phot}} < 0.6$ . When the density decreases, we see two different types of behaviour. For the Fisher\_GRIW1 selection, the percentage increases: this is because the mean redshift also increases, but staying at  $\lesssim 0.8$  implies that more galaxies are included in the  $0.6 \leq z_{\text{phot}} \leq 1.0$  range. However, we see a different behaviour for the three other selections, where the percentage flattens or starts to reverse when going to low densities. This is explained by the fact that the selections start to include galaxies at  $z_{\text{phot}} > 1.0$ . The reversal at low densities for the Fisher\_UGRIZW1 selection is due to the higher redshift of the selection and, for the Fisher\_GRZW1 selection, is due to a higher redshift and a broader redshift distribution. For low densities, the Fisher\_GRIW1<sup>OH</sup> selection has a slightly lower redshift than the Fisher\_UGRIZW1 selection and a slightly narrower distribution than the Fisher\_GRZW1 selection, which explains that we only observe a flattening and not a reversal.

Additionally, for the four Fisher discriminant selections, this percentage is fairly constant, within a few percentage points, for densities between ~150 deg<sup>-2</sup> and ~250 deg<sup>-2</sup>. Overall, this

means that increasing the selection density while lowering the threshold cut on the Fisher discriminant should still provide satisfactory results. We note that the percentage of galaxies with  $0.6 \leq z_{\text{phot}} \leq 1.0$  is higher than our computed efficiency with  $z_{\text{spec}}$ , because it does not require the additional criterion to have a reliable  $z_{\text{spec}}$  measurement. To quantitatively illustrate the impact of increasing the target density, we report in Table 5 the properties for the Fisher\_UGRIZW1 selection when the cut on the Fisher discriminant is set to have a target density of 210 deg<sup>-2</sup> ( $1.209 < X_{FI}$ ) over our ~50 deg<sup>2</sup> test area.

## 6. Fisher\_UGRIZW1 and Fisher\_GRIW1 selection stacked properties

We have shown in the previous section that the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections successfully select galaxies in the  $0.6 \leq z_{\text{spec}} \leq 1.0$  range with [OII] emission, thus permitting ~70% of the selection to be in the desired redshift range with a reliable  $z_{\text{spec}}$  measurement in the 1h observation with the BOSS spectrograph. Although allowing a reliable redshift measurement, the typical individual spectra are noisy (see Figure 3), which prevents us from visualising or measuring the typical features of the selected galaxies. Stacking the data allows us to significantly increase the signal-to-noise

		CC_UGRI	CC_UgrizW1	CC_griW1
<b>Photometric data (<math>\sim 50 \text{ deg}^2</math>)</b>				
(L1)	Density ( $\text{deg}^{-2}$ )	183	179	183
(L2)	eBOSS/LRG overlap ( $\text{deg}^{-2}$ )	2.4	1.8	2.6
(L3)	median( $z_{\text{phot}}$ )	0.75	0.77	0.77
(L4)	$z_{\text{phot}}$ peak width	0.13	0.12	0.12
(L5)	% with $0.6 \leq z_{\text{phot}} \leq 1.0$	76%	81%	83%
(L6)	mean( $z_{\text{phot}}$ ) ( $0.6 \leq z_{\text{phot}} \leq 1.0$ )	0.78	0.78	0.78
<b>VIPERS (<math>0.6 \leq z_{\text{spec}} \leq 1.0</math>)</b>				
(L7)	N selected galaxies	536	666	668
(L8)	$f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$	64%	76%	73%
<b>Plates eboss6-7 (<math>8.82 \text{ deg}^2</math>)</b>				
(L9)	N selected galaxies	1604	1670	1692
(L10)	Targeted	96%	(81%)	(87%)
(L11)	Unreliable $z_{\text{spec}}$	26%	(10%)	(13%)
(L12)	Median( $z_{\text{spec}}^{\text{reliable}}$ )	0.74	(0.77)	(0.76)
(L13)	$z_{\text{spec}}^{\text{reliable}}$ peak width	0.12	(0.12)	(0.12)
(L14)	Efficiency ( $0.6 < z_{\text{spec}}^{\text{reliable}} < 1.0$ )	59%	(76%)	(73%)
<b>Plates eboss6-7 <math>0.6 &lt; z_{\text{spec}}^{\text{reliable}} &lt; 1.0</math> only</b>				
(L15)	Mean( $z_{\text{spec}}^{\text{reliable}}$ )	0.77	(0.78)	(0.78)
(L16)	$f_{[\text{OII}]}^{\text{tot}} > 10^{-16} \text{ erg.cm}^{-1}.\text{s}^{-2}$	78%	(86%)	(85%)
(L17)	Expected $z_{\text{spec}}$ failure	1.1%	(0.5%)	(0.7%)

**Table 4.** Summary of the colour-colour selections properties. Lines are similar to Table 3

	Fisher_UGRIZW1
<b>Photometric data (<math>\sim 50 \text{ deg}^2</math>)</b>	
Density ( $\text{deg}^{-2}$ )	212
eBOSS/LRG overlap ( $\text{deg}^{-2}$ )	2.1
median( $z_{\text{phot}}$ )	0.77
$z_{\text{phot}}$ peak width	0.12
Percentage with $0.6 \leq z_{\text{phot}} \leq 1.0$	78%
mean( $z_{\text{phot}}$ ) ( $0.6 \leq z_{\text{phot}} \leq 1.0$ )	0.78
<b>VIPERS (<math>0.6 \leq z_{\text{spec}} \leq 1.0</math>)</b>	
N selected galaxies	671
$\log_{10}(f_{[\text{OII}]}^{\text{tot}}) > -16.0$	78%
<b>Plates eboss6-7 (<math>8.82 \text{ deg}^2</math>)</b>	
N selected galaxies	1902
Targeted	86%
Non-reliable $z_{\text{spec}}$	(12%)
median( $z_{\text{spec}}^{\text{reliable}}$ )	(0.77)
$z_{\text{spec}}^{\text{reliable}}$ peak width <sup>††</sup>	(0.12)
Efficiency ( $0.6 \leq z_{\text{spec}}^{\text{reliable}} \leq 1.0$ )	(69%)
<b>Plates eboss6-7, <math>0.6 \leq z_{\text{spec}}^{\text{reliable}} \leq 1.0</math> only</b>	
mean( $z_{\text{spec}}^{\text{reliable}}$ )	(0.78)
$\log_{10}(f_{[\text{OII}]}^{\text{tot}}) > -16.0$	(81%)
Expected $z_{\text{spec}}$ failure	(0.7%)

**Table 5.** Table illustrating the Fisher\_UGRIZW1 selection properties for a target density of  $210 \text{ deg}^{-2}$  ( $1.209 < X_{FI}$ ) over our  $\sim 50 \text{ deg}^2$  test area. Reported quantities are similar to those in Table 3.

ratio in the data, thereby enabling this visualisation or measurement of typical features of the selections that would not be visible or measurable in the individual data.

In this section, we take advantage of the unbiased, almost complete coverage of the Fisher\_UGRIZW1 and Fisher\_GRIW1

selections with the ten eboss6-7 test plates to study some global physical properties of those two selections through the use of stacked data. We stack the eboss6-7 spectra in Section 6.1 and the CFHTLS-Wide images in Section 6.2.

### 6.1. Stacked spectra

To illustrate the typical spectral features of the  $0.6 \leq z_{\text{spec}} \leq 1.0$  galaxies passing the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections, we display the median stacked spectra of the observed galaxies passing the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections and having  $0.6 \leq z_{\text{spec}} \leq 1.0$  (about 1100 galaxies per stack) in the top panel of Figure 13. We display the rest-frame wavelength range falling in the BOSS spectrograph from  $z_{\text{spec}} = 0.6$  to  $z_{\text{spec}} = 1.0$ . The stacked spectra have a  $S/N$  of  $\sim 5$  in this range. To enhance the differences between the two selections in the bottom panel of Figure 13, we display the median stacked spectra, removing objects that at the same time pass the two selections. The stack is thus done with  $\sim 300$  galaxies and has a  $S/N$  of  $\sim 3$ .

Firstly, the clear [OII] emission line in both stacked spectra confirms the ELG nature of the selections. Though both selections have comparable stacked spectra at first order, they nevertheless present small differences: the Fisher\_UGRIZW1 stacked spectrum has stronger emission lines, a smaller  $4000 \text{ \AA}$  break, and more emission in the near ultra-violet. All point to a more star-forming nature for the Fisher\_UGRIZW1 selection compared to the Fisher\_GRIW1 selection, in agreement with the analysis done in the Section 5.4. Furthermore, as expected, those differences are increased in the bottom panel of Figure 13. We recall that this stacking is only used for a qualitative visualisation of the typical spectral features of the  $0.6 \leq z_{\text{spec}} \leq 1.0$  galaxies passing the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections. A more thorough analysis of those



stacked spectra is presented in Paper I.

## 6.2. Stacked CFHTLS-Wide image

We took advantage of the eboss6-7 plates being within the CFHTLS-wide survey W1 field to do a simple morphological analysis for the two Fisher\_UGRIZW1 and Fisher\_GRIW1 selections. The CFHTLS-Wide images are about three magnitudes deeper than the SDSS and have better resolution and seeing (pixel scale of  $0.187'' \cdot \text{pix}^{-1}$  and seeing of  $0.7''$ - $0.8''$ ). Galaxy surface brightness distribution can be modelled with a Sérsic (1968) profile  $I(r) = I_e \times \exp\{-\kappa[(r/r_e)^{1/n_{\text{ser}}} - 1]\}$ , where  $I(r)$  is the surface brightness at  $r$ , and  $I_e$  is the surface brightness at the effective radius  $r_e$ , which is the radius which encloses half of the emitted light. The Sérsic index  $n_{\text{ser}}$  translates the shape of the profile, with a higher value corresponding to a profile more peaked at the centre and with larger wings:  $n_{\text{ser}} = 4$  corresponds to a de Vaucouleurs (1948) profile, which is typical of early-type galaxies, while  $n_{\text{ser}} = 1$  corresponds to an exponential profile, typical of late-type galaxies. Wuyts et al. (2011) have shown that, for  $0.1 < z_{\text{spec}} < 2.5$ , typical passive galaxies have  $n_{\text{ser}} \sim 4$ , while typical star-forming galaxies have  $n_{\text{ser}} \sim 1$ -2.

We used the CFHTLS  $i$ -band images and restricted ourselves to the eboss6-7 galaxies passing the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections with  $0.7 \leq z_{\text{spec}} \leq 0.8$  to mitigate the effect of redshift on the angular size of the galaxies. In this redshift range, which corresponds to the peak of the  $z_{\text{spec}}$  distribution for the two considered selections, the  $i$  band probes the rest frame  $4100 \text{ \AA}$ - $4400 \text{ \AA}$ . We obtain similar results if we use the CFHTLS  $r$ -band images, which probe the  $3400 \text{ \AA}$ - $3700 \text{ \AA}$  rest-frame at  $0.7 \leq z_{\text{spec}} \leq 0.8$ . For each stamp used in the median stacked image, we masked neighbouring objects beforehand, subtracted the sky, and scaled the galaxy fluxes to a normalised absolute magnitude. For both selections, we created two median stacked images, using the eboss6-7 galaxies with  $0.7 \leq z_{\text{spec}} \leq 0.8$ : 1) We used all the observed objects passing the selection ( $\sim 370$  galaxies); 2) we used only objects passing one selection but not the other ( $\sim 90$  galaxies). We fitted the surface brightness distribution with the GALFIT software (v3.0.5: Peng et al. 2010). During the fit, we set the axis ratio to 1 and the position angle to 0; besides this, we used spectroscopic stars ( $18 \leq i_{\text{AB}} \leq 21$ ) in the eboss6-7 plates area to create a point-spread function (PSF) stamp.

We present the stacked images and their radial profile in the Figure 14. We also report in this figure the estimated  $r_e$  and  $n_{\text{ser}}$ , along with their uncertainty computed via a thousand bootstrap realisations for each case. The stacked galaxies have small sizes, though clearly resolved: the surface brightness profile extends significantly farther than the PSF full-width-half-maximum. Regarding the surface brightness profile shape, we observe that both selections have a Sérsic index of  $\sim 1.3$ - $1.4$ , typical of star-forming galaxies (top panel). Interestingly, we see that the Fisher\_GRIW1 selection galaxies have a slightly higher Sérsic index, which is consistent with the trend seen in previous sections toward the Fisher\_GRIW1 selection forming fewer stars than the Fisher\_UGRIZW1 selection. This trend is more significant on the stacked images using galaxies belonging to only one of the selections (bottom panel). In addition, the Fisher\_GRIW1 selection galaxies also have sizes slightly larger than the Fisher\_UGRIZW1 selection galaxies ones. Finally, we note that the Fisher\_UGRIZW1 selection galaxies tend to have

relatively more flux in a  $2''$  aperture – corresponding to the BOSS spectrograph fibre diameter – than the Fisher\_GRIW1 selection, though this effect is minor. For the fitted parameters corresponding to the bottom panel of Figure 14, 90% of the total flux is included in a  $2''$  aperture for the Fisher\_UGRIZW1 selection galaxies, versus 84% for the Fisher\_GRIW1 selection galaxies.

## 7. Conclusions

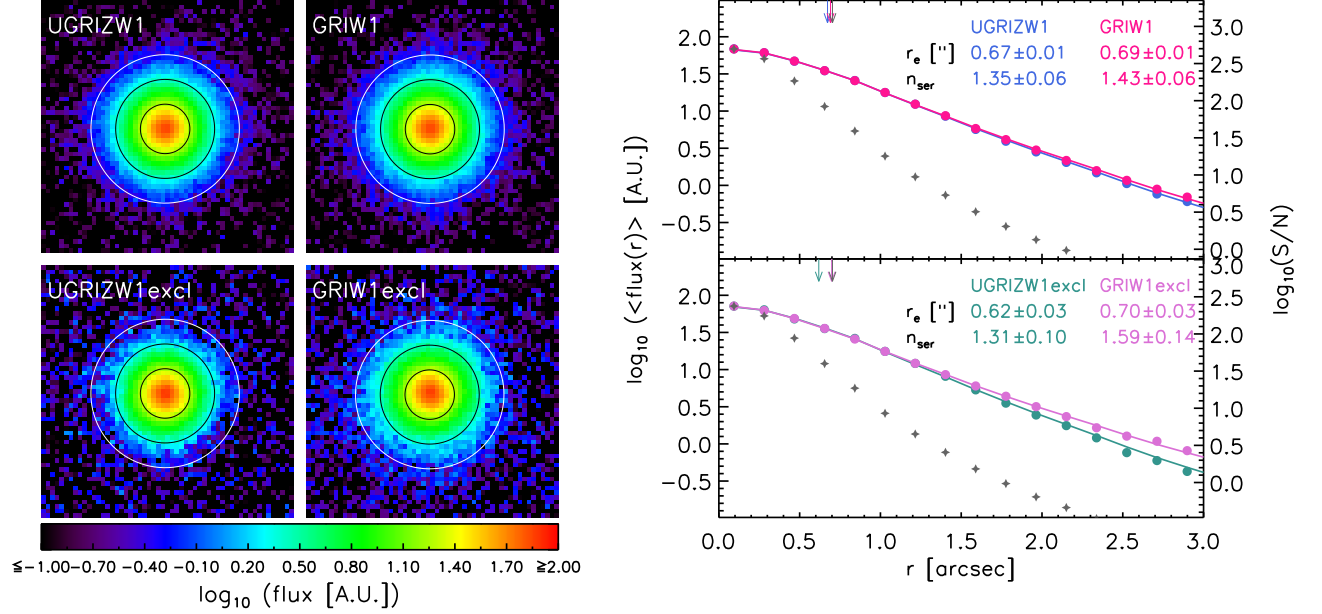
We have studied possible  $z \sim 0.8$  ELG selection schemes in preparation of the eBOSS/ELG survey. The initial eBOSS/ELG requirements are to select  $180 \text{ deg}^{-2}$  SDSS galaxies, 70% of which have a reliable  $z_{\text{spec}}$  measurement in a  $\sim 1\text{h}$  exposure observation with the BOSS spectrograph,  $0.6 \leq z_{\text{spec}} \leq 1.0$ , and a catastrophic failure rate  $\leq 1\%$  in this redshift range. Our selection schemes are based on the Fisher discriminant approach, which consists in computing the Fisher discriminant, a linear combination of colours defined from a spectroscopic training sample, and a simple selection with cuts on magnitudes and on this Fisher discriminant. This type of selection is simple and has the advantage of being flexible, since the density can be adjusted by modifying the Fisher discriminant threshold.

We studied the use of different photometric surveys: SCUSS/ $u$ +SDSS/ $grz$ +WISE/W1, SDSS/ $gri$ +WISE/W1, and SDSS/ $gr$ +DECaLS/ $z$ +WISE/W1. We quantified the properties of our selections in terms of redshift, [OII] emission, and efficiency, using dedicated eBOSS/ELG test plates and public photometric and spectroscopic data. We did a parallel analysis of colour-colour selections and showed, on the one hand, that the W1-band is crucial in improving the efficiency and, on the other hand, that the Fisher selections are competitive with colour-colour selections.

The Fisher\_UGRIZW1 selection meets the eBOSS/ELG redshift requirements. It has a median redshift of 0.78, and 68% of the selection has  $0.6 \leq z_{\text{spec}} \leq 1.0$ ; among those 68% galaxies,  $\sim 80\%$  have a significant [OII] emission ( $f_{[\text{OII}]}^{\text{tot}} \geq 10^{-16} \text{ erg.cm}^{-2} \cdot \text{s}^{-1} \cdot \text{\AA}^{-1}$ ), and the catastrophic  $z_{\text{spec}}$  measurement is expected to be 0.6% with the current instrumental setup and pipeline. The Fisher\_GRIW1 selection also meets the eBOSS/ELG redshift requirements with a 69% efficiency in the  $0.6 \leq z_{\text{spec}} \leq 1.0$  range and an expected catastrophic  $z_{\text{spec}}$  measurement of 1.1% in this redshift range. This selection has, on average, less [OII] emission than the Fisher\_UGRIZW1 selection. Training the Fisher method with  $z_{\text{spec}}$  and [OII] flux (Fisher\_UGRIZW1<sup>OII</sup> selection) allows us to increase the [OII] emission of the selection, but at the cost of a slightly lower mean redshift. The Fisher\_GRZW selection using the DECaLS/ $z$ -band seems to provide an acceptable alternative if set to a  $180 \text{ deg}^{-2}$  target density. Finally, we show that the density can be increased while keeping a reasonably high number of galaxies with  $0.6 \leq z_{\text{phot}} \leq 1.0$ .

In addition, we also studied the properties of the stacked spectra and stacked CFHTLS-Wide images for the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections. Those stacked data present typical features of star-forming galaxies and also indicate that the Fisher\_UGRIZW1 selection tends to favour more star-forming galaxies than the Fisher\_GRIW1 selection.

For the two most efficient selections, the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections, the homogeneity over the SGC along with the possible systematic dependence on various quan-



**Fig. 14.** CFHTLS-Wide *i*-band stacked images and structural parameters for the eboss6-7 plates galaxies with  $0.7 \leq z_{spec} \leq 0.8$  passing the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections. We labelled UGRIZW1 and GRIW1 when the stacking was done using all galaxies with  $0.7 \leq z_{spec} \leq 0.8$  and passing the selections ( $\sim 370$  galaxies per stack). We labelled UGRIZW1excl and GRIW1excl when the stacking was done using only galaxies with  $0.7 \leq z_{spec} \leq 0.8$  and not belonging to the intersection between Fisher\_UGRIZW1 and Fisher\_GRIW1 selections ( $\sim 90$  galaxies per stack), to enhance the differences. *Left panel:* stacked images. Each displayed image has a  $10''$  width and the circles have a  $1''$ ,  $2''$  and  $3''$  radius. *Right panel:* radial profile computed from the stacked images. Dots and diamonds represent the data, solid lines represent the fitted profile, corresponding to the reported values. The grey stars indicate the PSF profile. The arrows on the top x-axis represent the fitted  $r_e$  and the PSF FWHM.

tities, is studied in Paper III. Paper III also presents the catalogue release over the SGC for those two selections.

To conclude, this Fisher discriminant approach can be used more generally if one desires to select a galaxy population with desired properties with multi-band photometry. The method is simple and offers flexibility, the only requirement being the use of a spectroscopic training sample. Future massive spectroscopic surveys, such as DESI, 4MOST<sup>11</sup>, or the Prime Focus Spectrograph (PFS; Sugai et al. 2012), will provide large samples that are particularly well-suited to the Fisher discriminant approach.

**Acknowledgements.** AR acknowledges funding from the P2IO LabEx (ANR-10-LABX-0038) in the framework "Investissements d'Avenir" (ANR-11-IDEX-0003-01) managed by the French National Research Agency (ANR). JC acknowledges financial support from MINECO (Spain) under project number AYA2012-31101. JPK and TD acknowledge support from the ERC advanced grant LIDA. This study is based on data from SDSS-III (full text acknowledgement: <http://www.sdss3.org/collaboration/boiler-plate.php>) and is done in the context of SDSS-IV (full text acknowledgement: <http://www.sdss.org/collaboration/#OfficialSDSSAcknowledgment>). The SCUSS is funded by the Main Direction Program of Knowledge Innovation of Chinese Academy of Sciences (No. KJCX2-EW-T06). It is also an international cooperative project between National Astronomical Observatories, Chinese Academy of Sciences and Steward Observatory, University of Arizona, USA. Technical support and observational assistances of the Bok telescope are provided by Steward Observatory. The project is managed by the National Astronomical Observatory of China and Shanghai Astronomical Observatory. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, the Jet Propulsion Laboratory/California Institute of Technology, and NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. WISE and NEOWISE are funded by the National Aeronautics and Space Administration. This work is based on observations

obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. This research uses data from the VIMOS VLT Deep Survey, obtained from the VVDS database operated by Cesam, Laboratoire d'Astrophysique de Marseille, France. This paper uses data from the VIMOS Public Extragalactic Redshift Survey (VIPERS). VIPERS was performed with the ESO Very Large Telescope under the "Large Programme" 182.A- 0886. The participating institutions and funding agencies are listed at <http://vipers.inaf.it>. We thank the anonymous referee for his report, which helped us improve the clarity of the paper.

## References

- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12
- Anderson, L., Aubourg, E., Bailey, S., et al. 2012, *MNRAS*, 427, 3435
- Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
- Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Comparat, J., Delubac, T., Jouvel, S., et al. 2015a, *A&A*, submitted [arXiv:1509.05045]
- Comparat, J., Richard, J., Kneib, J.-P., et al. 2015b, *A&A*, 575, A40
- Coupon, J., Ilbert, O., Kilbinger, M., et al. 2009, *A&A*, 500, 981
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
- de Vaucouleurs, G. 1948, *Annales d'Astrophysique*, 11, 247
- Delubac, T., Bautista, J. E., Busca, N. G., et al. 2014, *ArXiv e-prints*
- Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, *MNRAS*, 401, 1429
- Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, *AJ*, 122, 2267
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
- Fisher, R. A. 1936, *Annals of Eugenics*, 7, 179
- Font-Ribera, A., Kirkby, D., Busca, N., et al. 2014, *J. Cosmology Astropart. Phys.*, 5, 27
- Franzetti, P., Garilli, B., Guzzo, L., Marchetti, A., & Scodreggio, M. 2014, *A&A*, submitted [arXiv:1410.0465]

<sup>11</sup> <https://www.4most.eu/>

Garilli, B., Guzzo, L., Scodeggio, M., et al. 2014, A&A, 562, A23  
 Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108  
 Gwyn, S. D. J. 2012, AJ, 143, 38  
 Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355  
 Hopkins, A. M. & Beacom, J. F. 2006, ApJ, 651, 142  
 Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841  
 Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236  
 Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55  
 Ilbert, O., Tresse, L., Zucca, E., et al. 2005, A&A, 439, 863  
 Jovel, S., Delubac, T., Comparat, J., et al. 2015, MNRAS, submitted [arXiv:1509.07121]  
 Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 54  
 Lang, D., Hogg, D. W., & Schlegel, D. J. 2014, AJ, submitted [arXiv:1410.7397]  
 Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14  
 Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 4841, 1670  
 Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845  
 Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218  
 Lilly, S. J., Le Fèvre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1  
 Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106  
 Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5  
 Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, AJ, 139, 2097  
 Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148  
 Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525  
 Seo, H.-J., Ho, S., White, M., et al. 2012, ApJ, 761, 13  
 Sersic, J. L. 1968, Atlas de galaxias australes (Observatorio Astronomico, Cordoba, Argentina)  
 Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32  
 Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810  
 Sugai, H., Karoji, H., Takato, N., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 0  
 The Dark Energy Survey Collaboration. 2005, arXiv:astro-ph/0510346  
 van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 709, 1018  
 Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868  
 Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, ApJ, 742, 96  
 York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579  
 Zou, H., Jiang, Z., Zhou, X., et al. 2015, AJ, 150, 104

observe that the photometric redshifts are slightly biased high ( $bias \sim +0.01$ , for all magnitudes), the bias becoming significant for  $z_{phot} \gtrsim 1.1$ . The scatter is reasonable ( $\sigma_{outl.rej.} \sim 0.04$ ), as is the outlier rate (5-10%). Those results qualitatively agree with previous studies (Ilbert et al. 2006; Hildebrandt et al. 2012).

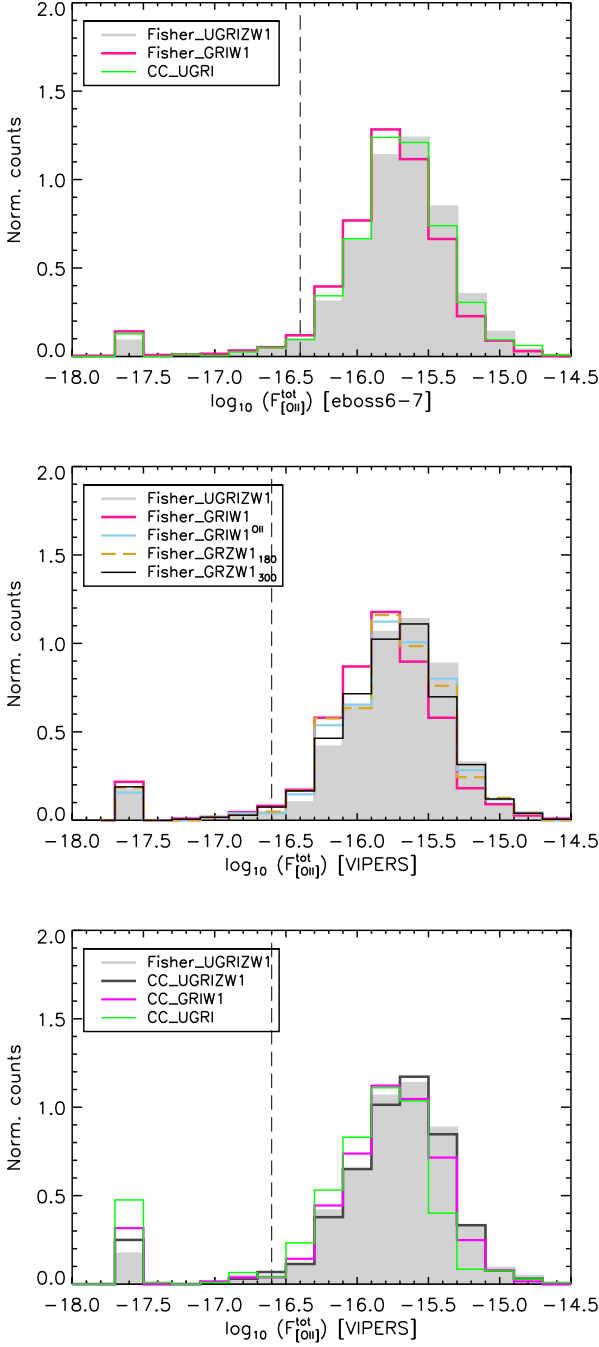
## Appendix A: CFHTLS photometric redshift reliability

In Section 5.4, we make use of the CFHTLS-Wide photometric redshifts (T0007 release<sup>12</sup>; Ilbert et al. 2006; Coupon et al. 2009) to estimate the redshift distribution of our selection schemes. Those photometric redshifts have been proven to be of very good quality up to  $i < 22.5$  (bias below 1%, scatter of  $\sim 0.04$ , and less than 4% outliers). Nevertheless, those statistics have been computed for magnitude-limited samples, whereas the galaxies under study in this paper are mainly star-forming galaxies. Owing to the lack of features (weak 4000 Å break, power law spectrum), this class of galaxies is well-known for having slightly less accurate photometric redshift, which results in a higher outlier rate (e.g. Ilbert et al. 2006; Hildebrandt et al. 2012).

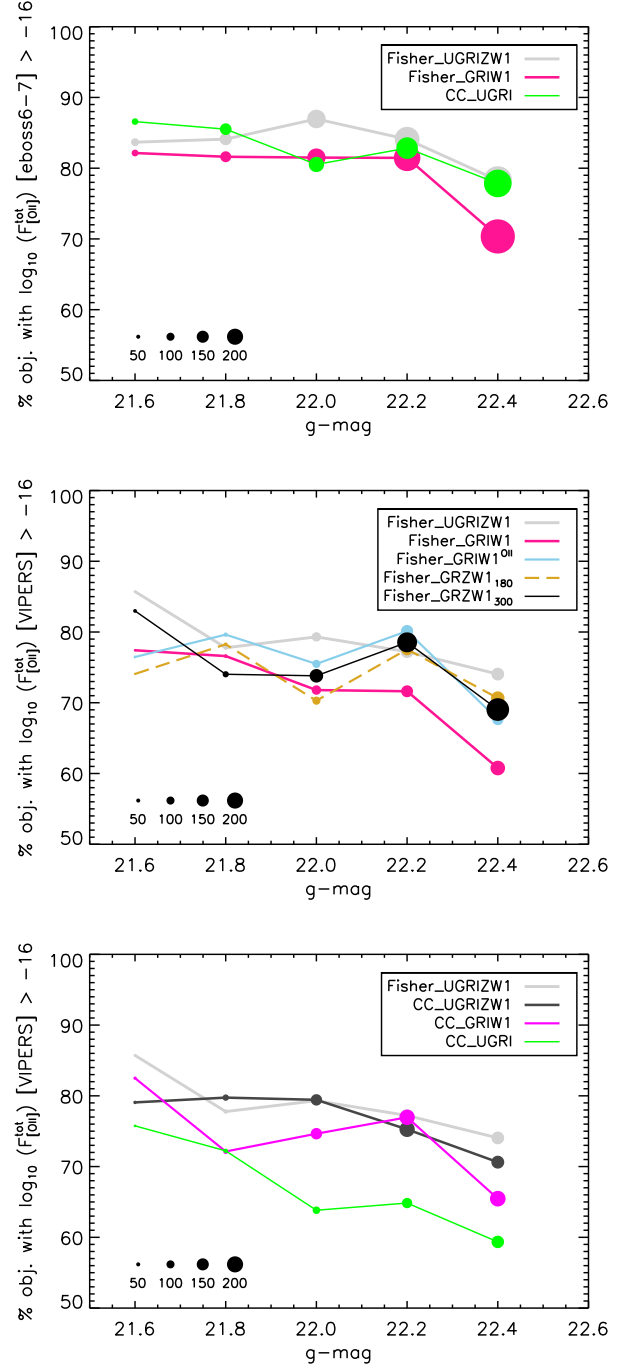
Using the eboss6-7 test plates, we demonstrate in Figure A.1 the reliability of those photometric redshifts for ELGs up to redshift  $\sim 1$ . We selected eboss6-7 galaxies with a secure spectroscopic redshift and a [OII] total luminosity  $L_{[OII]}^{tot}$  greater than  $10^{41} \text{ erg.s}^{-1}$  (3712 galaxies with  $z_{spec} = 0.80 \pm 0.17$ ). For each object in our spectroscopic sample, we calculated  $\Delta z = \frac{z_{phot} - z_{spec}}{1 + z_{spec}}$  and classify it as an outlier if  $|\Delta z| > 0.15$ . For each binned subsample, we report  $bias$ : the median value of  $\Delta z$ ;  $outl.$ : the percentage of outliers; and  $\sigma_{outl.rej.}$ : the standard deviation of  $\Delta z$  when outliers have been excluded. These quantities are used to facilitate comparison with other works. As mentioned in Hildebrandt et al. (2012), the outlier definition is arbitrary. We

<sup>12</sup> [http://terapix.iap.fr/rubrique.php?id\\_rubrique=267](http://terapix.iap.fr/rubrique.php?id_rubrique=267)

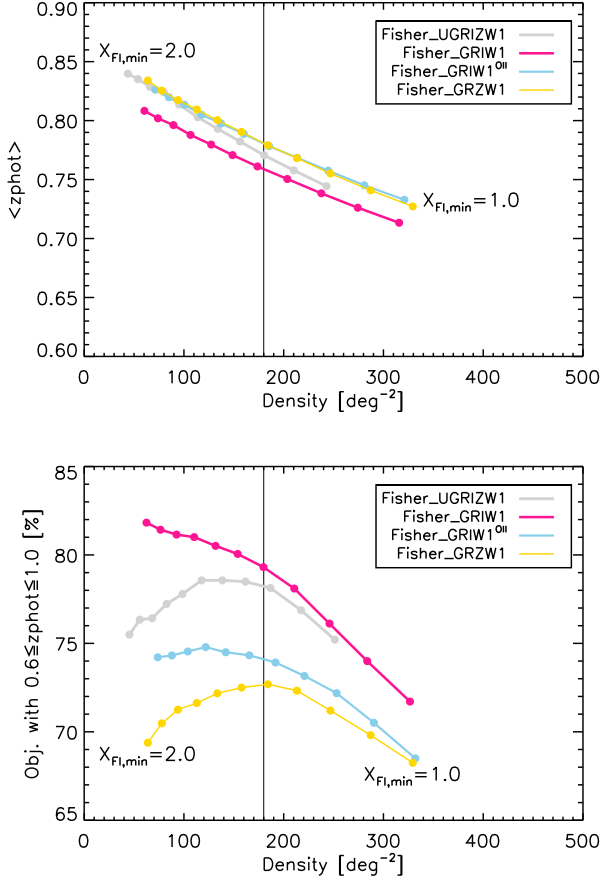




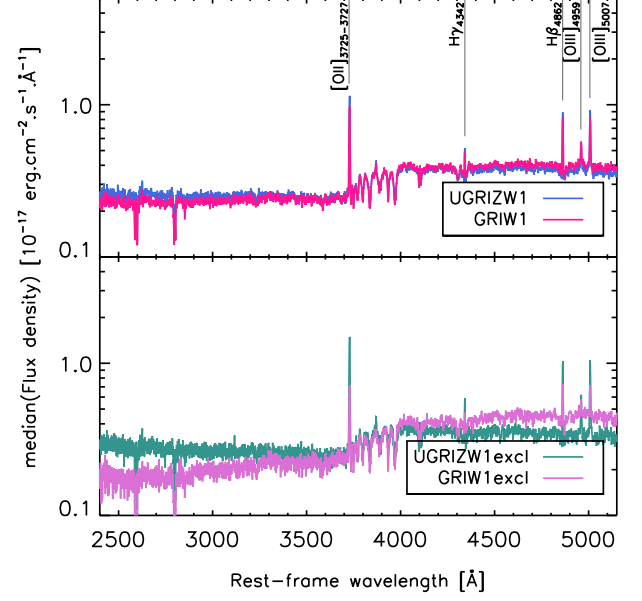
**Fig. 10.** Total [OII] flux distributions for the five Fisher and the CC\_UGRI selections. The vertical dashed line represents the approximate  $F_{\text{[OII]}}^{\text{tot}}$  of objects with  $S/N = 3$ . Objects with unreliable [OII] flux measurement are represented at  $f_{\text{[OII]}}^{\text{tot}} = -17.6$ . We consider here only those galaxies with a reliable  $z_{\text{spec}}$  measurement. *Top panel:* eboss6-7 galaxies. *Middle panel:* VIPERS galaxies for the Fisher selections. *Bottom panel:* VIPERS galaxies for the colour-colour selections. We report the Fisher\_UGRIZW1 selection to facilitate the comparison.



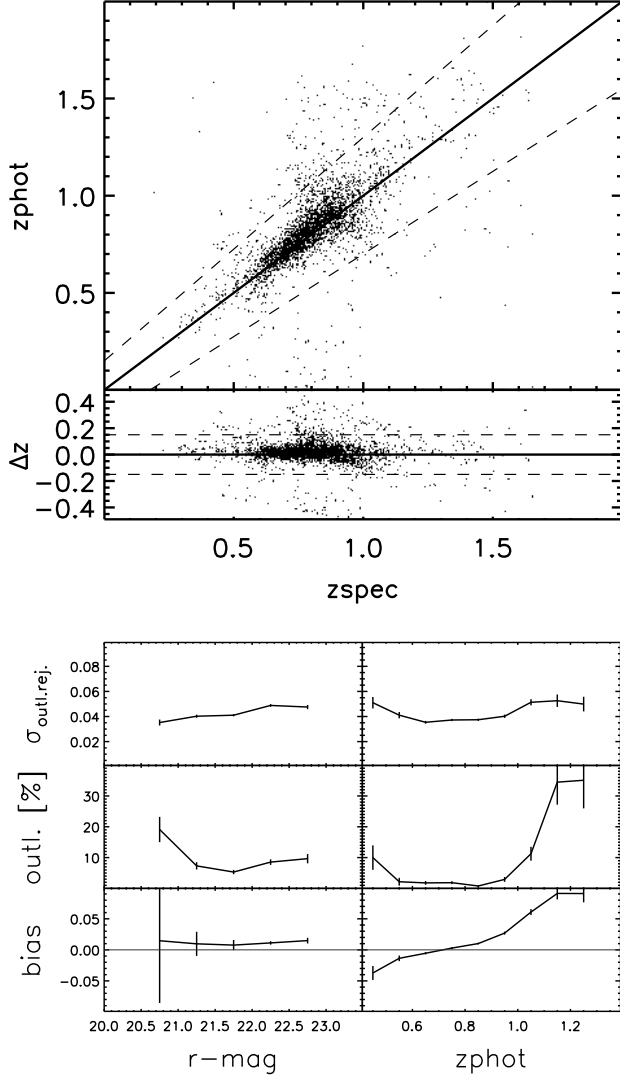
**Fig. 11.** Total [OII] flux distributions vs. the  $g$ -band magnitude for the five Fisher and the CC\_UGRI selections. We consider here only galaxies with a reliable  $z_{\text{spec}}$  measurement. The dot size scales with the number of objects entering the bin. *Top panel:* eboss6-7 galaxies. *Middle panel:* VIPERS galaxies for the Fisher selections. *Bottom panel:* VIPERS galaxies for the colour-colour selections. We report the Fisher\_UGRIZW1 selection to facilitate the comparison.



**Fig. 12.** Selection dependence on  $X_{FI,min}$ , the Fisher discriminant threshold cut.  $X_{FI,min}$  varies from 2.0 (low densities, left) to 1.0 (high densities, right) with a step of 0.1. The sky region considered here is our  $\sim 50 \text{ deg}^2$  test area within the CFHTLS W1 field. *Top panel:* mean  $z_{\text{phot}}$  of the selection. *Bottom panel:* percentage of galaxies with  $0.6 \leq z_{\text{phot}} \leq 1.0$  estimated with the CFHTLS  $z_{\text{phot}}$ . The thin vertical line illustrates the  $180 \text{ deg}^2$  density used to set the Fisher discriminant threshold cut in Section 5.4.



**Fig. 13.** Stacked spectra from the eboss6-7 plates for the Fisher\_UGRIZW1 and Fisher\_GRIW1 selections. *Top panel:* the stacking (average  $S/N$  of  $\sim 5$ ) is done using all galaxies passing the selections ( $\sim 1100$  galaxies per stack). *Bottom panel:* the stacking (average  $S/N$  of  $\sim 3$ ) is done using only galaxies not belonging to the intersection between Fisher\_UGRIZW1 and Fisher\_GRIW1 selections ( $\sim 300$  galaxies per stack), to enhance the differences.



**Fig. A.1.** Reliability of the CFHTLS photometric redshifts for  $z \sim 0.8$  ELGs, using the eboss6-7 observed galaxies with a secure  $z_{\text{spec}}$  and  $L_{[\text{OII}]}^{\text{tot}} > 10^{41} \text{ erg.s}^{-1}$ . *Top panel:*  $z_{\text{phot}}$  vs.  $z_{\text{spec}}$ . *Bottom panel:*  $z_{\text{phot}}$  statistics, as a function of magnitude (*left*) and redshift (*right*). We only report quantities for the bins where we have more than 50 galaxies, and error bars are calculated assuming a Poissonian distribution.